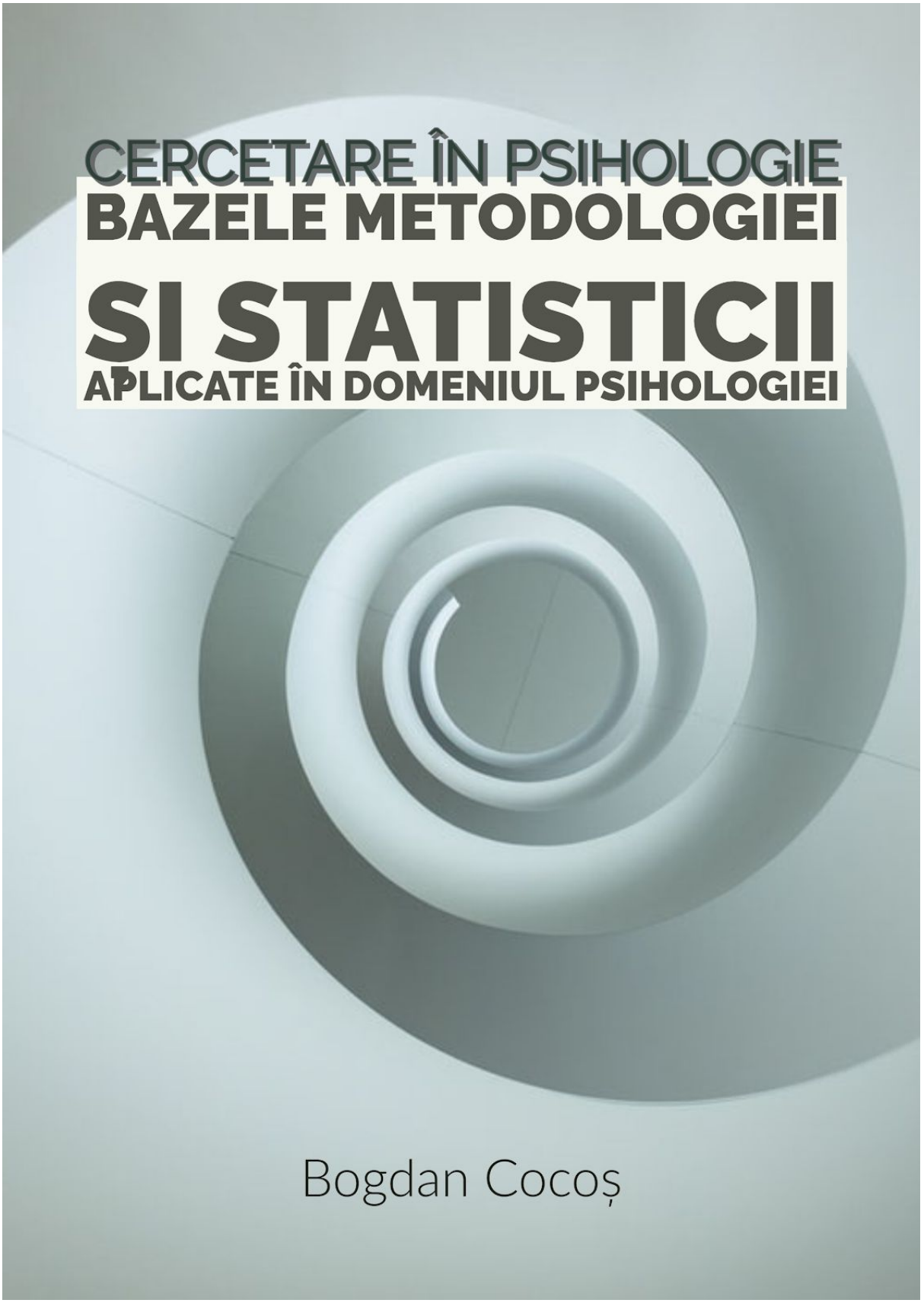
# CERCETARE ÎN PSIHOLOGIE
# BAZELE METODOLOGIEI
# ȘI STATISTICII
## APLICATE ÎN DOMENIUL PSIHOLOGIEI



Bogdan Cocoș

# Cercetare în Psihologie: Bazele Metodologiei și Statisticii Aplicate în Domeniul Psihologiei

Teorie și Exemple


## Bogdan Cocoș

Master Sănătate Ocupațională și Performanța Resursei Umane, Facultatea de Psihologie și Științele Educației, Universitatea din București

# Cuprins

**Eşantionarea** **7**
Tipuri de Selectare 7
Selectarea Aleatorie 7
Selectarea Nealeatorie 7
Tipuri de Alocare 8
Alocarea Aleatorie 8
Alocarea Nealeatorie 8
Diferenţa dintre Selectare şi Alocare 8

**Validitatea** **10**
Validitatea Testelor 10
Validitatea de Construct 10
Validitatea de Criteriu 10
Validitatea Studiilor 10
Validitatea Concluziei Statistice 10
Validitatea Internă 11
Validitatea Externă 11

**Eficienţă, Eficacitate şi Efectivitate** **13**
Eficienţa (Efficacy) 13
Eficacitatea (Effectiveness) 13
Efectivitatea (Efficiency) 13

**Design-uri de Cercetare Experimentale** **16**
Concepte de Bază 16
Trei Tipuri de Design-uri în Block-uri 17
Design Complet Aleatoriu 17
Design Aleatoriu în Block-uri 18
Design de Tip Pătrat Latin 20
Design-uri Factoriale 21
Design Factorial Complet Aleatoriu 21
Modele Alternative 22
Model de Regresie 22
Cell Means Model 22
Design Factorial Aleatoriu în Block-uri 22
Design-uri Factoriale cu Confundate/ Covariate 24
Fundamente Teoretice 24
Design Factorial de Tip Split-Plot 25
Design-uri Factoriale Fracţionare 26

Design-uri Ierarhice 26
Design-uri Ierarhice cu Unul/ Două Tratamente Grupate 26
Design Ierarhic cu Tratamente Încrucişate şi Grupate 27
Design-uri Experimentale cu O Covariată 28
Analiza de Covarianţă (ANCOVA) 28

**Design-uri de Cercetare Cvasi-Experimentale 32**
Experiment pe Serii de Timp (Time-Series Experiment) 32
Design pe Eşantioane Temporale Echivalente (Equivalent Time-Sample Design) 32
Design cu Materiale/ Instrumente Echivalente (Equivalent Materials Design) 32
Design Neechivalent cu Grup de Control (Nonequivalent Control Group Design) 33
Design-uri Contrabalansate (Counterbalanced Designs) 33
Design cu Eşantion Separat Pretest-Posttest (Separate Sample Pretest-Posttest Design) 33
Design cu Eşantion Separat Pretest-Posttest şi Grup de Control (Separate Sample Pretest-Posttest Control Group Design) 34
Design pe Serii Multiple de Timp (Multiple Time-Series Design) 34
Design-ul Ciclului Instituţional Recurent (Recurrent Institutional Cycle Design) 34
Analiza Regresie-Discontinuitate (Regression-Discontinuity Analysis) 35

**Design-uri de Cercetare Non-Experimentale 39**
Studiu Transversal (Cross-Sectional) 39
Definiţie şi Caracteristici 39
Avantaje şi Dezavantaje 39
Studiu de Tip Case-Control 39
Definiţie şi Caracteristici 39
Avantaje şi Dezavantaje 39
Studiu de Tip Înainte-După (Before-and-After) 40
Definiţie şi Caracteristici 40
Avantaje şi Dezavantaje 40
Studiu Retrospectiv de Tip Istoric cu Grup de Control (Historical Controls) 40
Definiţie şi Caracteristici 40
Avantaje şi Dezavantaje 40
Sondaje/ Chestionare 41
Definiţie şi Caracteristici 41
Avantaje şi Dezavantaje 41
Studiu de Caz de Tip Case Series 41
Definiţie şi Caracteristici 41
Avantaje şi Dezavantaje 41
Studiu de Caz de Tip Case Report 42
Definiţie şi Caracteristici 42
Avantaje şi Dezavantaje 42

**Alte Design-uri de Cercetare 45**
Studii Corelaţionale 45

Definiţie şi Caracteristici 45
Limitări 45
Meta-Analize 45
Definiţie şi Caracteristici 45
Limitări 45
Revizuiri Sistematice (Systematic Reviews) 46
Definiţie şi Caracteristici 46
Limitări 46
Studii Cost-Beneficiu (Cost-Benefit) 46
Definiţie şi Caracteristici 46
Limitări 46
Studii Cost-Eficacitate (Cost-Effectiveness) 46
Definiţie şi Caracteristici 46
Limitări 46

**Ierarhia Evidenţelor** **49**
Accepţia Clasică 49
Scurte Definiţii 49
Case Series şi Case Reports 49
Case Control Studies 49
Cohort Studies 50
Randomized Controlled Trials (RCTs) 50
Systematic Review şi Meta-Analysis 50
Ediţia Revizuită 50

**Fidelitatea** **52**
Definiţie şi Tipuri de Fidelitate 52
Fidelitatea Inter-Evaluator 52
Definiţie şi Caracteristici 52
Statistici Uzuale 52
Procentajul de Acord 52
Kappa-ul lui Cohen 52
Corelaţia Produs-Moment 53
Coeficientul de Corelaţie Intra-Clasă 53
Fidelitatea Test-Retest/ Intra-Evaluator 53
Fidelitatea de Tip Forme Paralele/ Forme Echivalente 53
Fidelitatea de Tip Consistenţă Internă 54
Definiţie şi Caracteristici 54
Statistici Uzuale 54
α-ul lui Cronbach 54
Corelaţia Medie Inter-Itemi 54
Fidelitatea Split Half 54
Kuder-Richardson 20 (KR-20) 54

**Analiza de Putere** **56**
Tipuri de Erori Statistice 56
Definiţie şi Caracteristici 56
Factori ce Pot Influenţa Puterea Unui Test 57
Analiza Prospectivă (A Priori) 57
Analiza Retrospectivă (Post Hoc) 57

**Statistici Descriptive Univariate** **59**
Tendinţa Centrală a Distribuţiei 59
Media (Mean/ Average) 59
Mediana (Median) 59
Modul (Mode) 59
Dispersia Distribuţiei 60
Intervalul (Range) 60
Quartilele (Quartiles) 61
Abaterea Standard (Standard Deviation) 61
Varianţa (Variance) 62
Forma Distribuţiei 62
Skewness 62
Kurtosis 63
Histograme (Histograms) 63
Reprezentări Grafice de Tip Stem-and-Leaf 64

**Teste Statistice Elementare pentru Psihologie** **66**
Testul t pentru Un Singur Eşantion 66
Testul t pentru Eşantioane Perechi 66
Asumpţiile Testului 66
Mărimea Efectului 67
Testul t pentru Eşantioane Independente 67
Asumpţiile Testului 67
Mărimea Efectului 67
Analiza de Varianţă (ANOVA) Unifactorială 68
Asumpţiile Testului 68
Mărimea Efectului 68
Epsilon Squared 68
Eta Squared 68
Omega Squared 68
Corelaţie 68
Asumpţiile Testului 68
Regresie Liniară Simplă 70
Asumpţiile Testului 70
Mărimea Efectului 72

**Analiza de Moderare şi Analiza de Mediere** **74**
Analiza de Moderare 74
Definiţie şi Caracteristici 74
Asumpţii 74
Analiza de Mediere 75
Definiţie şi Caracteristici 75
Asumpţii 75
Mediere Totală vs. Mediere Parţială 77
**Semnificaţia Statistică** **79**
Definiţie 79
Accepţia Clasică 79
Accepţia Revizuită 79
Principiile Semnificaţiei Statistice 79
Semnificaţia Statistică, Mărimea Efectului şi Semnificaţia Clinică 79
Concluzii 79

# Metodologie

## Eșantionarea

### Tipuri de Selectare

- Aleatorie
    - Aleatorie simplă (*Simple random*).
    - Aleatorie stratificată (*Stratified random*).
    - Eșantionarea pe grupuri (*Cluster sampling*).
    - Eșantionarea sistematică (*Systematic sampling*).
    - Eșantionarea multistadială (*Multi-stage sampling*).
- Nealeatorie
    - Eșantionarea consecutivă/ pe cote (*Quota sampling*).
    - Eșantionarea de tip bulgăre de zăpadă (*Snowball sampling*).
    - Eșantionarea deliberată (*Judgement sampling*).
    - Eșantionarea de convenienţă (*Convenience sampling*).

#### Selectarea Aleatorie

Selectarea aleatorie simplă este un model de eșantionare în care *k* elemente distincte sunt selectate din *n* elemente posibile ale populaţiei, astfel încât fiecare combinaţie posibilă de *k* elemente are șansa egală de a fi eșantionul selectat (Thompson, 2012, în Meng, 2013).

Selectarea aleatorie stratificată constă în aplicarea eșantionării aleatorii simple pe subseturi omogene distincte și reprezentative pentru populaţie (Meng, 2013).

Eșantionarea pe grupuri este o tehnică de eșantionare în care toate elementele populaţiei sunt categorizate în grupuri mutual exclusive şi exhaustive (denumite "*clusters*") (Wakar & Poznyak, 2018).

Eșantionarea sistematică este o metodă de eșantionare ce aplică un interval constant (i.e., din *k* în *k*) în selectarea eșantioanelor dintr-un anumit cadru de eșantionare (i.e., din populaţia *N*, se extrage eșantionul *n*, urmând ca, din eșantionul *n* să fie selectat fiecare al *k*-lea element aleatoriu *r*) (Cox, 2008).

Eșantionarea multistadială este o metodă de eșantionare ce conţine 2 sau mai multe stadii de eșantionare aleatorie simplă bazată pe structura ierarhică a grupurilor natural formate în interiorul populaţiei (e.g., școli, gospodării) (Sedgwick, 2015).

#### Selectarea Nealeatorie

Eșantionarea consecutivă/ pe cote este o metodă de eșantionare în care eșantioanele sunt selectate în baza probabilităţii ca o anumită variabilă din populaţie să fie distribuită proporţional (e.g., proporţia de eșantioane dintr-o categorie are proporţia teoretică a populaţiei) (Rukmana, 2014).

Eșantionarea de tip bulgăre de zăpadă este un tip de eșantionare de convenienţă în care respondenţii prezenţi în cadrul unui studiu recrutează viitori respondenţi, dintre cunoștinţe (e.g., rude, prieteni) (Naderifar, Goli, & Ghaljaei, 2017).

Eşantionarea deliberată este o metodă de eşantionare în care unităţile selectate pentru un studiu sunt incluse în baza judecăţii profesionale a cercetătorului (Maul, 2018).

Eşantionarea de convenienţă este o metodă de eşantionare în care respondenţii sunt selectaţi în baza faptului că reprezintă surse "convenience" de date pentru cercetători (i.e., prezintă disponibilitate) (Battaglia, 2008).

# Tipuri de Alocare

- Aleatorie
- Nealeatorie

## Alocarea Aleatorie

Alocarea aleatorie reprezintă procesul prin care cercetătorii selectează aleatoriu indivizi din eşantionul format, cu scopul de a-i aloca unor condiţii specifice. Fiecare respondent are o probabilitate specifică de a fi alocat într-o anumită condiţie. Condiţiile (denumite "grupuri") reprezintă niveluri ale variabilei independente (Salkind, 2010a).

## Alocarea Nealeatorie

Alocarea nealeatorie reprezintă alocarea respondenţilor selectaţi din eşantion în baza anumitor caracteristici. Astfel, participanţii nu au şanse egale de a fi plasaţi în grupuri.

# Diferența dintre Selectare și Alocare

Selectarea presupune selectarea indivizilor dintr-o populaţie. În contrast, alocarea reprezintă alocarea indivizilor selectaţi în grupuri (e.g., experimental, control), fie aleatoriu, fie nealeatoriu.

Selectarea este o componentă a validităţii externe (i.e., priveşte generalizabilitatea rezultatelor), în timp ce alocarea reprezintă o componentă a validităţii interne.

## Resurse Online - Eşantionarea

1. Distinguish between participant selection and participant assignment and tell the specific type of validity associated with each
2. Difference between Random Selection and Random Assignment
3. Eşantionarea 1 Eşantionul
4. Sampling Methods | Types and Techniques Explained
5. Sampling | Research Methods Knowledge Base
6. Sampling in Statistics: Different Sampling Methods, Types & Error
7. Types of sampling methods | Statistics (article)
8. RANDOM.ORG - True Random Number Service
9. http://www.youtube.com/watch?v=cTA5WWLhUp0

# Validitatea

## Validitatea Testelor

- Validitatea de construct
    - Validitatea de conţinut.
    - Validitatea de faţadă.
- Validitatea de criteriu
    - Validitatea concurentă.
    - Validitatea predictivă.

### Validitatea de Construct

Validitatea de conţinut reprezintă măsura în care itemii, afirmaţiile sau întrebările unui test sunt reprezentativi/ reprezentative pentru constructul măsurat de test (e.g., itemii ce fac referire la depresie sunt reprezentativi pentru măsurarea depresiei) (Salkind, 2010b).

Validitatea de faţadă reprezintă măsura în care testul pare că măsoară ceea ce îşi propune să măsoare (i.e., testul pare bine făcut, în viziunea respondenţilor) (Johnson, 2013).

### Validitatea de Criteriu

Validitatea concurentă reprezintă gradul de acord dintre două evaluări independente (e.g., măsurarea depresiei cu instrumentul A este similară cu măsurarea depresiei cu instrumentul B) (Adams, Cervantes, Jang, & Dixon, 2014).

Validitatea predictivă reprezintă gradul de acord dintre rezultatele obţinute cu ajutorul unui test nou şi rezultatele obţinute cu ajutorul unui test acreditat (i.e., Gold Standard) (Multani, Saini, Kaur, & Sharma, 2014).

## Validitatea Studiilor

- Validitatea concluziei statistice.
- Validitatea internă.
- Validitatea externă.

### Validitatea Concluziei Statistice

Validitatea concluziei statistice poate fi văzută drept o parte mică a unui proiect de cercetare sau a unui studiu al cărui ţel constă în comunicarea unei concluzii şi în luarea unei decizii în baza analizelor statistice efectuate.

Majoritatea erorilor privind validitatea concluziei statistice apar din cauza faptului că există o înţelegere precară a procedurilor statistice, rezultând într-o utilizare eronată a acestora (García-Pérez, 2012).

## Validitatea Internă

Validitatea internă reprezintă măsura în care rezultatele observate reprezintă adevăruri privind populaţia-ţintă şi nu efecte cauzate de erori metodologice (Patino & Ferreira, 2018).

Dintre factorii ce pot afecta validitatea internă, amintim: istoricul, maturizarea, efectele testării, instrumentele utilizate, regresia statistică, erorile de selectare, mortalitatea, ordinea temporală a cauzalităţii, difuzia tratamentelor, rivalitatea compensatorie, demoralizarea, compensarea, efectele resimţite de cercetători şi efectele resimţite de participanţi. Pentru mai multe detalii, în privinţa acestor factori, a se consulta primul link listat în cadrul secţiunii Resurse Online, a acestui capitol.

## Validitatea Externă

Validitatea externă reprezintă măsura în care efectele observate în cadrul studiului pot fi observate şi în afara acestuia (i.e., generalizabilitatea unui studiu) (Khorsan & Crawford, 2014).

Dintre factorii ce pot afecta validitatea externă, amintim: erorile de selectare, constructele măsurate, metodele şi variabilele confundate, diferenţa dintre "lumea reală" (i.e., semnificaţia practică) şi "lumea experimentală" (i.e., semnificaţia statistică), istoricul şi maturizarea. Pentru mai multe detalii, în privinţa acestor factori, a se consulta cel de-al doilea link listat în cadrul secţiunii Resurse, a acestui capitol.

Un tip excepţional de validitate externă este validitatea ecologică. Aceasta reprezintă congruenţa sau asocierea dintre testele derulate într-un cadru specific (e.g., laborator, clinică) şi sarcinile de lucru cotidiene (Johansson & Wahlin, 1998).

## Resurse Online - Validitatea

1. Internal validity | Lærd Dissertation
2. External validity | Lærd Dissertation
3. Validity
4. The 4 Types of Validity | Explained with Easy Examples
5. Reliability and Validity of Measurement – Research Methods in Psychology
6. RESEARCH METHODS: PLANNING: VALIDITY
7. Research Methodology: Validity
8. Two Criteria for Good Measurements in Research: Validity and Reliability

# Eficiență, Eficacitate și Efectivitate

## Eficienţa (*Efficacy*)

Eficienţa poate fi definită ca abilitatea unei intervenţii de a obţine rezultatul vizat, în condiţii considerate a fi ideale (Crespo-Facorro et al., 2017).

Pentru a înţelege mai bine conceptul eficienţei, vor fi discutate, în continuare, caracteristicile acestuia.

În acest sens, consecinţele primare ale studiilor de eficienţă privesc reducerea simptomelor, schimbări comportamentale sau optimizări ale abilităţilor. Strategia tipic utilizată este cea de studiu clinic controlat (*RCT*, *Randomized Controlled Trial*). De asemenea, studiile de eficienţă se pot concentra pe mecanismele acţiunii intervenţiei. Studiile tipice de eficienţă iau în calcul criterii restrânse de producere a unui eşantion omogen (i.e., asumă faptul că respondenţii reprezintă populaţia-ţintă, fără un control prealabil al asumpţiei, rezultând, astfel, o validitate externă limitată) (Hargreaves, Catalano, Hu, & Cuffel, 1998). Un exemplu de utilizare a conceptului de eficienţă poate fi observat în articolul lui Swart şi Apsche (2014).

## Eficacitatea (*Effectiveness*)

Conceptul de eficacitate poate fi înţeles ca măsura în care efectul vizat este obţinut, în condiţii de practică rutinieră clinică (Crespo-Facorro et al., 2017).

Pentru a înţelege mai bine conceptul eficacităţii, vor fi discutate, în continuare, caracteristicile acestuia.

În acest sens, studiile de eficacitate pot fi găsite şi sub denumirile de "*real-world studies*" sau "*pragmatic trials*". Aceste studii se bazează pe criterii mai largi de excluziune, permiţând o mai mare varietate de pacienţi, condiţii de tratament şi niveluri ale psihopatologiilor asociate. Totodată, se bazează pe consecinţe privind calitatea vieţii şi preferinţele pentru instrumente de autoraportare sau globale. Au o mai mare validitate externă, accentuându-se potenţialul de generalizare a rezultatelor obţinute cu ajutorul acestor studii. Cu toate acestea, validitatea lor internă este limitată (Möller, 2011). Un exemplu de utilizare a conceptului de eficacitate poate fi observat în articolul lui Carter şi Fairburn (1998).

## Efectivitatea (*Efficiency*)

Conceptul de efectivitate nu are o traducere propriu-zisă în limba română. În acest sens, însăşi denumirea de "efectivitate" va fi utilizată drept o operaţionalizare a conceptului. În ciuda limitărilor lingvistice, efectivitatea reprezintă valoarea relativă a unei intervenţii, comparativ cu costul acesteia pentru individ sau societate (Crespo-Facorro et al., 2017).

Pentru a înţelege mai bine conceptul efectivităţii, vor fi discutate, în continuare, caracteristicile acestuia.

În acest sens, studiile de efectivitate reprezintă tentative de măsurare cât mai obiectivă a valorilor relative ale mai multor variabile (denumite "*estimators*") (Fisher, 1921, în Organisation for Economic Cooperation and Development, 2002). În general, studiile de efectivitate sunt utilizate pentru compararea intervenţiilor, în cadrul design-urilor experimentale. Astfel, un design

experimental ar putea fi mai efectiv decât un altul, în cazul în care dă dovadă de aceeaşi precizie, cu mai puţine resurse temporale sau financiare (Organisation for Economic Cooperation and Development, 2002). Un exemplu de utilizare a conceptului de efectivitate poate fi observat în articolul lui Yang şi Tsiatis (2001).

## Sumar - Eficienţă, Eficacitate şi Efectivitate

- Eficienţa - Rezultat vizat, condiţii ideale.
- Eficacitatea - Măsura eficienţei.
- Efectivitatea - Valoarea relativă temporal sau financiar a unei intervenţii, comparativ cu o alta.

## Bibliografie Selectivă - Eficienţă, Eficacitate şi Efectivitate


Carter, J. C., & Fairburn, C. G. (1998). Cognitive–behavioral self-help for binge eating disorder: A controlled effectiveness study. *Journal of Consulting and Clinical Psychology*, *66*(4), 616–623. https://doi.org/10.1037/0022-006X.66.4.616

Swart, J., & Apsche, J. (2014). A comparative treatment efficacy study of conventional therapy and mode deactivation therapy (MDT) for adolescents with conduct disorders, mixed personality disorders, and experiences of childhood trauma. *International Journal of Behavioral Consultation and Therapy*, *9*(1), 23-29. http://doi.org/10.1037/h0101011

Yang, L., & Tsiatis, A. A. (2001). Efficiency study of estimators for a treatment effect in a pretest–posttest trial. *The American Statistician*, *55*(4), 314-321. https://doi.org/10.1198/000313001753272466


## Resurse Online - Eficienţă, Eficacitate şi Efectivitate

1. Efficacy vs. Effectiveness vs. Efficiency – Healthcare Economist
2. Efficacy, Effectiveness and Efficiency in the Health Care: The Need for an Agreement to Clarify its Meaning
3. What is the difference between efficacy and effectiveness with respect to clinical trials?
4. Definitions: Efficacy, Effectiveness, Efficiency | Delivering Digital Drugs (D3)
5. The Difference Between Effectiveness and Efficiency Explained
6. Effectiveness vs. Efficiency vs. Efficacy

# Design-uri de Cercetare Experimentale

Înainte de a începe discuţia despre design-urile de cercetare experimentale, va fi listată, în continuare, o clasificare a acestora (Kirk, 2003, p. 11).

**TABLE 1.1 Classification of Experimental Designs**

| Experimental Design | Abbreviated Designation[a] |
|---|---|
| I. Systematic Designs (selected examples). | |
| 1. Beavan's chessboard design. | |
| 2. Beavan's half-drill strip design. | |
| 3. Diagonal square design. | |
| 4. Knut Vik square design. | |
| II. Randomized Designs With One Treatment. | |
| A. Experimental units randomly assigned to treatment levels. | |
| 1. Completely randomized design. | CR-$p$ |
| B. Experimental units assigned to relatively homogeneous blocks or groups prior to random assignment. | |
| 1. Balanced incomplete block design. | BIB-$p$ |
| 2. Cross-over design. | CO-$p$ |
| 3. Generalized randomized block design. | GRB-$p$ |
| 4. Graeco-Latin square design. | GLS-$p$ |
| 5. Hyper-Graeco-Latin square design. | HGLS-$p$ |
| 6. Latin square design. | LS-$p$ |
| 7. Lattice balanced incomplete block design. | LBIB-$p$ |
| 8. Lattice partially balanced incomplete block design. | LPBIB-$p$ |
| 9. Lattice unbalanced incomplete block design. | LUBIB-$p$ |
| 10. Partially balanced incomplete block design. | PBIB-$p$ |
| 11. Randomized block design. | RB-$p$ |
| 12. Youden square design. | YBIB-$p$ |
| III. Randomized Designs With Two or More Treatments. | |
| A. Factorial designs: designs in which all treatments are crossed. | |
| 1. Designs without confounding. | |
| a. Completely randomized factorial design. | CRF-$pq$ |
| b. Generalized randomized block factorial design. | GRBF-$pq$ |
| c. Randomized block factorial design. | RBF-$pq$ |
| 2. Design with group-treatment confounding. | |
| a. Split-plot factorial design. | SPF-$p \cdot q$ |
| 3. Designs with group-interaction confounding. | |
| a. Latin square confounded factorial design. | LSCF-$p^k$ |
| b. Randomized block completely confounded factorial design. | RBCF-$p^k$ |
| c. Randomized block partially confounded factorial design. | RBPF-$p^k$ |
| 4. Designs with treatment-interaction confounding. | |
| a. Completely randomized fractional factorial design. | CRFF-$p^{k-i}$ |
| b. Graeco-Latin square fractional factorial design. | GLSFF-$p^k$ |
| c. Latin square fractional factorial design. | LSFF-$p^k$ |
| d. Randomized block fractional factorial design. | RBFF-$p^{k-i}$ |
| B. Hierarchical designs: designs in which one or more treatments are nested. | |
| 1. Designs with complete nesting. | |
| a. Completely randomized hierarchical design. | CRH-$pq(A)$ |
| b. Randomized block hierarchical design. | RBH-$pq(A)$ |
| 2. Designs with partial nesting. | |
| a. Completely randomized partial hierarchical design. | CRPH-$pq(A)r$ |
| b. Randomized block partial hierarchical design. | RBPH-$pq(A)r$ |
| c. Split-plot partial hierarchical design. | SPH-$p \cdot qr(B)$ |
| IV. Randomized Designs With One or More Covariates. | |
| A. Designs that include a covariate have the letters AC added to the abbreviated designation as in the following examples. | |
| 1. Completely randomized analysis of covariance design. | CRAC-$p$ |
| 2. Completely randomized factorial analysis of covariance design. | CRFAC-$pq$ |
| 3. Latin square analysis of covariance design. | LSAC-$p$ |
| 4. Randomized block analysis of covariance design. | RBAC-$p$ |
| 5. Split-plot factorial analysis of covariance design. | SPFAC-$p \cdot q$ |
| V. Miscellaneous Designs (select examples). | |
| 1. Solomon four-group design. | |
| 2. Interrupted time-series design. | |

[a]The abbreviated designations are discussed later.

## Concepte de Bază

Design-urile experimentale sunt interesate de o analiză robustă a naturii. În acest sens, cercetătorii pot lua decizii informate privind ipotezele din ştiinţă şi tehnologie (i.e., prin compararea variabilităţii dintre participanţii care au primit intervenţii diferite cu variabilitatea participanţilor care au primit intervenţii similare) (Kirk, 2003).

Astfel, demersurile majoritar utilizate în descrierea design-urilor experimentale pot fi descrise în 5 paşi (Kirk, 2003):

1. Formularea ipotezelor statistice relevante pentru o ipoteză ştiinţifică.
2. Determinarea condiţiilor experimentale (i.e., variabile independente) ce vor fi manipulate, a măsurătorilor (i.e., variabile dependente) ce vor fi înregistrate şi a condiţiilor externe (i.e., variabile confundate/ zgomot) ce trebuie să fie controlate.
3. Specificarea numărului recomandat de participanţi şi a populaţiei din cadrul căreia aceştia sunt eşantionaţi.
4. Specificarea procedurii de alocare a participanţilor în condiţiile experimentale.
5. Determinarea analizelor statistice potrivite, ce vor fi derulate.

Sumar, există trei principii ale design-urilor experimentale bune (Kirk, 2003):

1. Controlul variabilei dependente.
2. Alocarea aleatorie a participanţilor în grupuri.
3. Includerea covariatei (i.e., variabilă-zgomot) ca un factor experimental.

# Trei Tipuri de Design-uri în Block-uri

## Design Complet Aleatoriu

Pentru schematizarea design-ului complet aleatoriu, vor fi folosite planurile generale utilizate de Kirk (2003, p. 5), respectiv un exemplu de utilizare a design-ului (da Silva, Fernandes, Colen, & Sampaio, 2017), după cum urmează.

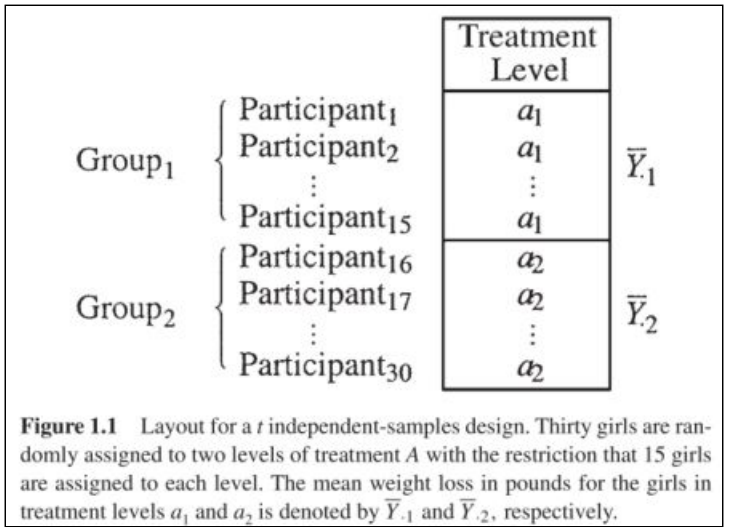



**Figure 1.1** Layout for a *t* independent-samples design. Thirty girls are randomly assigned to two levels of treatment *A* with the restriction that 15 girls are assigned to each level. The mean weight loss in pounds for the girls in treatment levels $a_1$ and $a_2$ is denoted by $\overline{Y}_{\cdot 1}$ and $\overline{Y}_{\cdot 2}$, respectively.

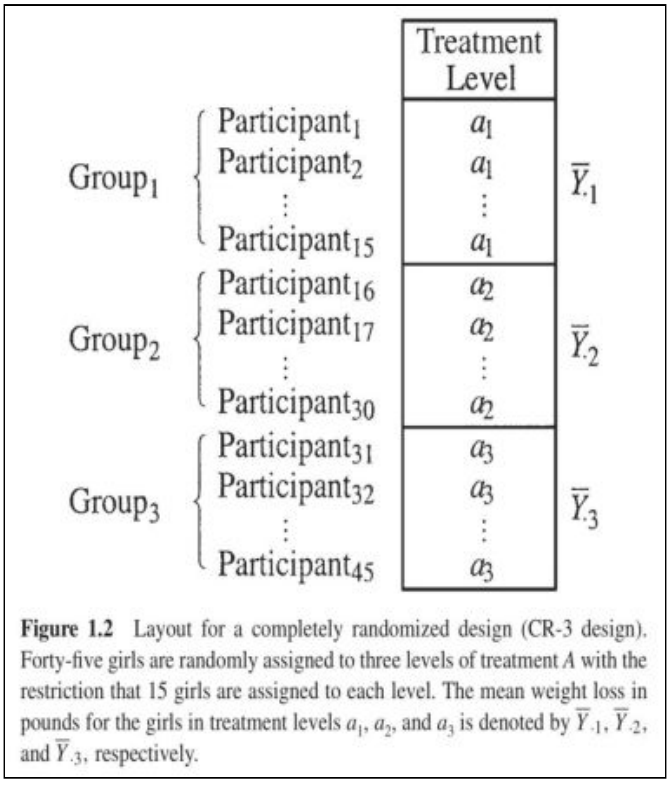



**Figure 1.2** Layout for a completely randomized design (CR-3 design). Forty-five girls are randomly assigned to three levels of treatment *A* with the restriction that 15 girls are assigned to each level. The mean weight loss in pounds for the girls in treatment levels $a_1$, $a_2$, and $a_3$ is denoted by $\overline{Y}_{\cdot 1}$, $\overline{Y}_{\cdot 2}$, and $\overline{Y}_{\cdot 3}$, respectively.

## Design Aleatoriu în Block-uri

Kirk (2003) delimitează o serie de paşi caracteristici design-urilor aleatorii în block-uri:

1. Observarea fiecărui participant pentru fiecare nivel/ tip de intervenţie, în experiment (i.e., măsurători repetate).
2. *Participant matching* - Formarea seturilor de participanţi similari din punctul de vedere al asocierii dintre intervenţie şi variabila dependentă.
3. Obţinerea de gemeni sau codescendenţi (i.e., pentru a se asigura controlul caracteristicilor genetice similare ale participanţilor).
4. Obţinerea participanţilor potriviţi prin selecţie mutuală (e.g., soţ-soţie, parteneri de afaceri, etc.).

Pentru o înţelegere aplicată a conceptului de design aleatoriu în block-uri, se va utiliza schema propusă de Kirk (2003, p. 7), după cum urmează.

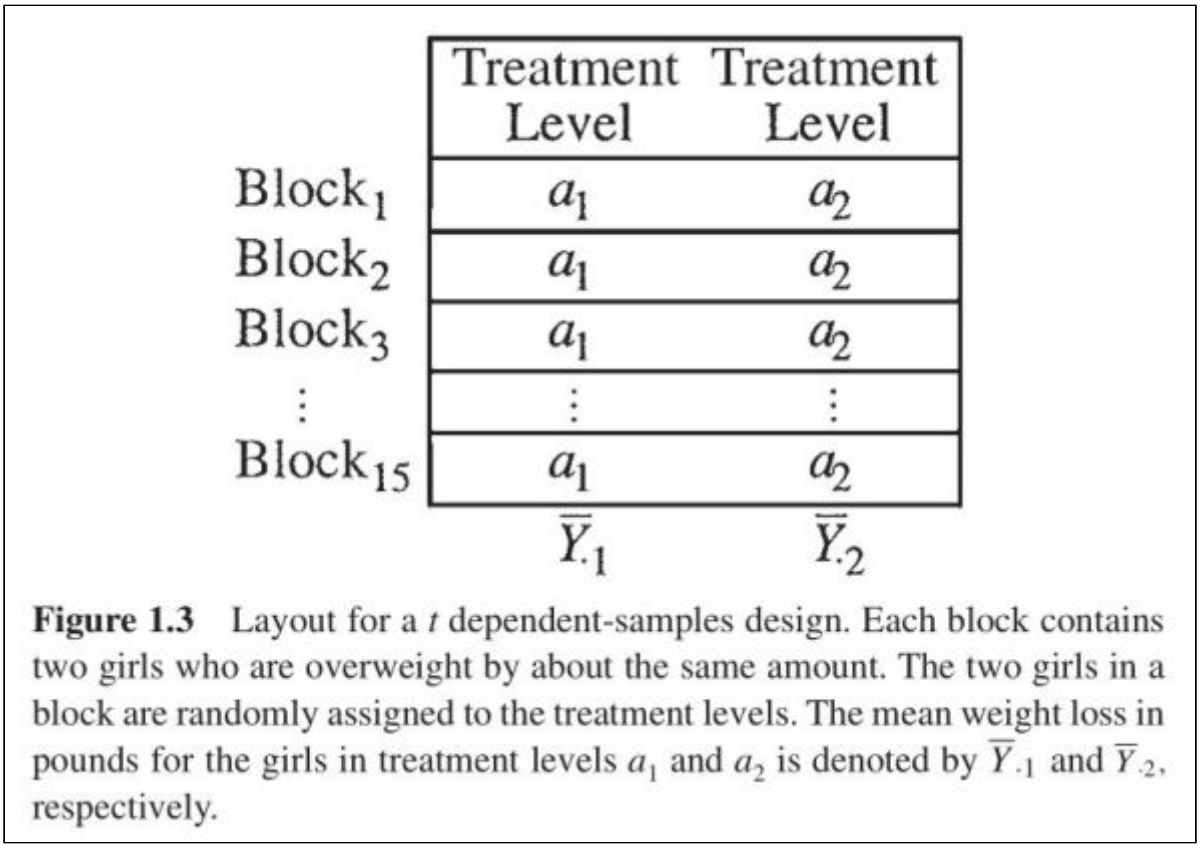



**Figure 1.3** Layout for a *t* dependent-samples design. Each block contains two girls who are overweight by about the same amount. The two girls in a block are randomly assigned to the treatment levels. The mean weight loss in pounds for the girls in treatment levels $a_1$ and $a_2$ is denoted by $\overline{Y}_{.1}$ and $\overline{Y}_{.2}$, respectively.

Consensul metodologic privind notarea acestui tip de design este RB-*p*, unde:

- RB = "Randomized Block".
- *p* = Numărul de niveluri ale intervenţiei.

Rezultatele unui experiment cu măsurători repetate pot fi generalizate la nivelul populaţiei de participanţi ce au fost expuşi la toate nivelurile intervenţiei. În contrast, rezultatele unui experiment cu participanţi-pereche pot fi generalizate la nivelul populaţiei de participanţi ce au fost expuşi la un singur nivel al intervenţiei (Kirk, 2003).

În acest sens, pot fi delimitate două tipuri de design-uri aleatorii în block-uri: (1) *subjects-by-treatment design*; şi (2) *subject-by-trials design*. Primul tip face referire la un design cu

măsurători repetate în care ordinea nivelurilor administrării intervenţiei sau a tratamentului este alocată aleatoriu, independent, pentru fiecare participant. Cel de-al doilea tip face referire la un design cu măsurători repetate în care ordinea nivelurilor administrării intervenţiei sau a tratamentului este aceeaşi pentru toţi participanţii (Kirk, 2003).

Pentru exemplificare, Kirk (2003, p. 8) propune un protocol de scădere a masei corporale, în funcţie de 3 niveluri posibile ale aceleiaşi intervenţii, după cum urmează.

| | Treatment Level | Treatment Level | Treatment Level | |
|---|---|---|---|---|
| $\text{Block}_1$ | $a_1$ | $a_2$ | $a_3$ | $\overline{Y}_{1\cdot}$ |
| $\text{Block}_2$ | $a_1$ | $a_2$ | $a_3$ | $\overline{Y}_{2\cdot}$ |
| $\text{Block}_3$ | $a_1$ | $a_2$ | $a_3$ | $\overline{Y}_{3\cdot}$ |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| $\text{Block}_{15}$ | $a_1$ | $a_2$ | $a_3$ | $\overline{Y}_{15\cdot}$ |
| | $\overline{Y}_{\cdot 1}$ | $\overline{Y}_{\cdot 2}$ | $\overline{Y}_{\cdot 3}$ | |

**Figure 1.4** Layout for a randomized block design (RB-3 design). Each block contains three girls who are overweight by about the same amount. The three girls in a block are randomly assigned to the treatment levels. The mean weight loss in pounds for the girls in treatment levels $a_1$, $a_2$, and $a_3$ is denoted by $\overline{Y}_{\cdot 1}$, $\overline{Y}_{\cdot 2}$, and $\overline{Y}_{\cdot 3}$, respectively. The mean weight loss for the girls in $\text{Block}_1$, $\text{Block}_2$, . . . , $\text{Block}_{15}$ is denoted by $\overline{Y}_{1\cdot}$, $\overline{Y}_{2\cdot}$, . . . , $\overline{Y}_{15\cdot}$, respectively.

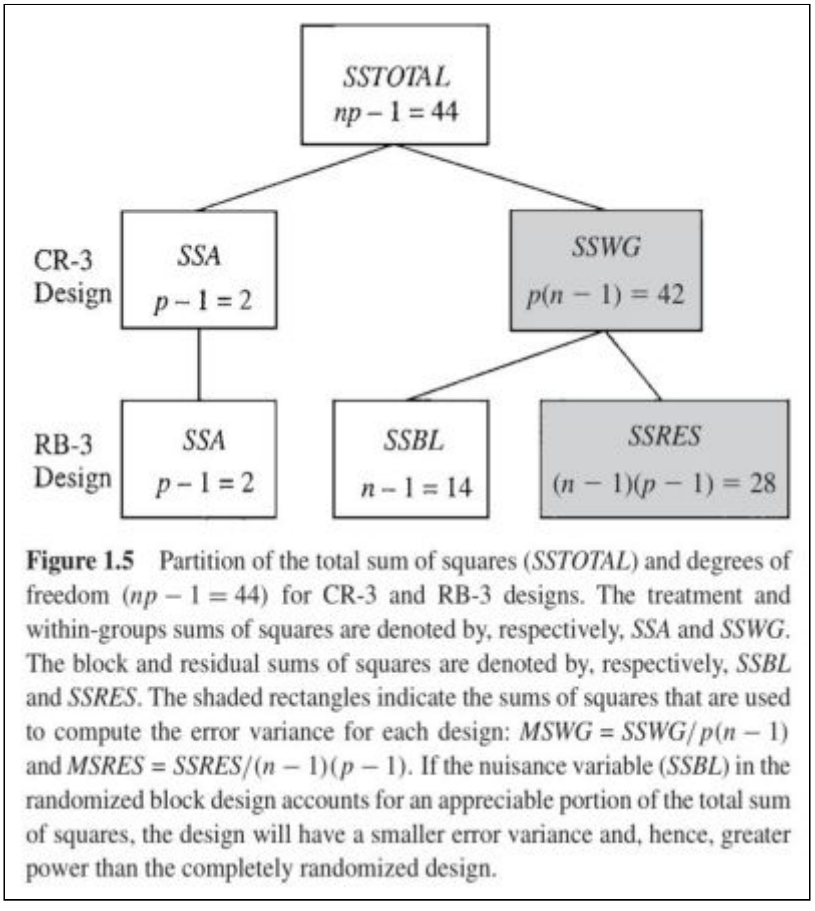



**Figure 1.5** Partition of the total sum of squares (*SSTOTAL*) and degrees of freedom ($np - 1 = 44$) for CR-3 and RB-3 designs. The treatment and within-groups sums of squares are denoted by, respectively, *SSA* and *SSWG*. The block and residual sums of squares are denoted by, respectively, *SSBL* and *SSRES*. The shaded rectangles indicate the sums of squares that are used to compute the error variance for each design: $MSWG = SSWG/p(n-1)$ and $MSRES = SSRES/(n-1)(p-1)$. If the nuisance variable (*SSBL*) in the randomized block design accounts for an appreciable portion of the total sum of squares, the design will have a smaller error variance and, hence, greater power than the completely randomized design.

Ca o lectură adiţională şi, bineînţeles, opţională, va fi propus şi un studiu (Chen et al., 2018) ce ţine de domeniul psihologiei (i.e., mai specific, neuropsihologie), pentru o înţelegere aprofundată a design-ului parcurs.

## Design de Tip Pătrat Latin

Numele acestui tip de design derivă de la un puzzle străvechi, ce viza în câte moduri pot fi aranjate literele într-o matrice pătratică (e.g., 3*3, n*n, etc.), astfel încât fiecare literă apare o singură dată în fiecare linie şi coloană (Kirk, 2003). Pentru analogie, ne putem duce cu gândul înspre jocul de *sudoku*.

Consensul metodologic privind notarea acestui tip de design este LS-*p*, unde:

- LS = "Latin Square".
- *p* = Numărul de niveluri ale tratamentului.

Un astfel de design îi permite cercetătorului să izoleze efectele a două variabile-zgomot. În acest sens, Kirk (2003, p. 9) propune o schemă generală a design-ului, după cum urmează.

| | $c_1$ | $c_2$ | $c_3$ |
|---|---|---|---|
| $b_1$ | $a_1$ | $a_2$ | $a_3$ |
| $b_2$ | $a_2$ | $a_3$ | $a_1$ |
| $b_3$ | $a_3$ | $a_1$ | $a_2$ |

**Figure 1.6** Three-by-three Latin square, where $a_j$ denotes one of the $j = 1, \ldots, p$ levels of treatment $A$; $b_k$ denotes one of the $k = 1, \ldots, p$ levels of nuisance variable $B$; and $c_l$ denotes one of the $l = 1, \ldots, p$ levels of nuisance variable $C$. Each level of treatment $A$ appears once in each row and once in each column as required for a Latin square.

| | | Treatment Combination | |
|---|---|---|---|
| $\text{Group}_1$ | $\text{Participant}_1$ ⋮ $\text{Participant}_5$ | $a_1b_1c_1$ ⋮ $a_1b_1c_1$ | $\overline{Y}_{\cdot 111}$ |
| $\text{Group}_2$ | $\text{Participant}_6$ ⋮ $\text{Participant}_{10}$ | $a_1b_2c_3$ ⋮ $a_1b_2c_3$ | $\overline{Y}_{\cdot 123}$ |
| $\text{Group}_3$ | $\text{Participant}_{11}$ ⋮ $\text{Participant}_{15}$ | $a_1b_3c_2$ ⋮ $a_1b_3c_2$ | $\overline{Y}_{\cdot 132}$ |
| $\text{Group}_4$ | $\text{Participant}_{16}$ ⋮ $\text{Participant}_{20}$ | $a_2b_1c_2$ ⋮ $a_2b_1c_2$ | $\overline{Y}_{\cdot 212}$ |
| ⋮ | ⋮ | ⋮ | ⋮ |
| $\text{Group}_9$ | $\text{Participant}_{41}$ ⋮ $\text{Participant}_{45}$ | $a_3b_3c_1$ ⋮ $a_3b_3c_1$ | $\overline{Y}_{\cdot 331}$ |

**Figure 1.7** Layout for a Latin square design (LS-3 design) that is based on the Latin square in Figure 1.6. Treatment $A$ represents three kinds of diets; nuisance variable $B$ represents amount by which the girls are overweight; and nuisance variable $C$ represents genetic predisposition to be overweight. The girls in $\text{Group}_1$, for example, received diet $a_1$, were less than fifteen pounds overweight ($b_1$), and neither parent had been overweight as a teenager ($c_1$). The mean weight loss in pounds for the girls in the nine groups is denoted by $\overline{Y}_{\cdot 111}, \overline{Y}_{\cdot 123}, \ldots, \overline{Y}_{\cdot 331}$.

După cum v-ați obișnuit deja, probabil, va fi citat, în continuare, un articol (Pine et al., 2016) în care veți putea observa utilizarea practică a acestui tip de design.

# Design-uri Factoriale

## Design Factorial Complet Aleatoriu

Schema generală a acestui tip de design (Kirk, 2003, p. 12) urmează astfel.

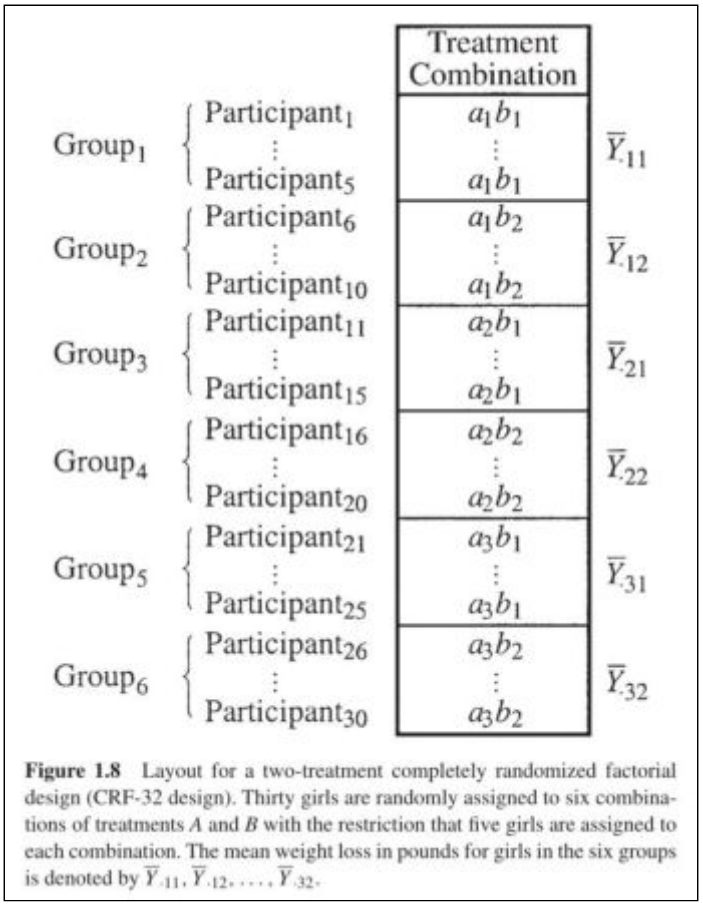



**Figure 1.8** Layout for a two-treatment completely randomized factorial design (CRF-32 design). Thirty girls are randomly assigned to six combinations of treatments *A* and *B* with the restriction that five girls are assigned to each combination. The mean weight loss in pounds for girls in the six groups is denoted by $\overline{Y}_{\cdot 11}, \overline{Y}_{\cdot 12}, \ldots, \overline{Y}_{\cdot 32}$.

Consensul metodologic privind notarea acestui tip de design este CRF-*pq*, unde:

- CRF = "Completely Randomized Factorial".
- *p* = Numărul de niveluri ale intervenţiei A.
- *q* = Numărul de niveluri ale intervenţiei B.

În acest punct, este necesară o discuţie despre semnificaţia statistică a rezultatelor obţinute în urma analizelor datelor colectate cu ajutorul design-urilor proaspăt învăţate. Firește, discuţia despre semnificaţia statistică a rezultatelor va fi întâlnită, în calitate de subcapitol independent, la sfârşitul capitolului dedicat statisticii.

Aşadar, este cunoscut faptul că există diferenţe între semnificaţia statistică şi semnificaţia practică a rezultatelor obţinute. Respingerea ipotezei de nul nu ne oferă prea multe informaţii, deoarece cunoaştem, deja, faptul că ipotezele formulate sunt false. Totodată, respingerea ipotezei de nul nu ne vorbește despre mărimile efectelor specifice tratamentelor sau despre importanța practică a

acestora. Astfel, cercetătorii ar trebui să se concentreze mai mult pe datele în sine, în detrimentul concentrării pe ce "spun" datele despre ipotezele ştiinţifice (Kirk, 2003). În acest sens, Hays (1963, în Kirk, 2003) a introdus o măsură a magnitudinii asocierilor dintre variabile, anume *Omega Squared* (i.e., înţeleasă ca proporţia varianţei de populaţie în variabila dependentă, de care este responsabil tratamentul). Urmând instrucţiunile oferite de Cohen (1988, pp. 284-288, în Kirk, 2003) pentru interpretarea indicelui *Omega Squared*, putem observa următoarele:

- .010 - Asociere slabă.
- .059 - Asociere medie.
- .138 - Asociere puternică.

Pentru o mai bună aplicare şi înţelegere a design-ului factorial complet aleatoriu, articolul lui Adam şi Hadi (2018) se poate dovedi util.

## Modele Alternative

În cazul în care numărul de condiţii (i.e., "celulele" unui model) nu este egal, unii cercetători utilizează anumite proceduri, pentru aproximarea testelor ipotezei de nul, după cum urmează: (1) estimarea observaţiilor-lipsă, sub asumpţia că tratamentele nu interacţionează; (2) eliminarea aleatorie de date, pentru a egala numărul condiţiilor; sau (3) utilizarea unei analize bazate pe calcule efectuate manual (i.e., *unweighted-means analysis*) (Kirk, 2003).

Pentru o înţelegere aprofundată a modelelor alternative, Kirk (2003) discută despre două tipuri de astfel de modele:

- Model de regresie.
- *Cell means model*.

### Model de Regresie

Pentru detalierea acestui tip de model, poate fi consultat subcapitolul dedicat analizei de regresie liniară, din cadrul acestei lucrări, sau paginile 35-44 ale materialului valabil pe OSF, la adresa: https://osf.io/tx82n/, din cadrul proiectului "Cercetare în Psihologie" (Cocoş, 2020).

### *Cell Means Model*

Acest tip de model evită problemele asociate cu supraparametrizarea (i.e., utilizarea unui număr excesiv de parametri, în analizele statistice). Astfel, media unei populaţii poate fi estimată pentru fiecare condiţie care conţine una sau mai multe măsurători (i.e., modelul este complet parametrizat). Acest tip de model nu impune o structură asupra analizei datelor (i.e., poate fi utilizat pentru testarea ipotezelor asociate cu orice combinaţie liniară dintre mediile condiţiilor populaţiei) (Kirk, 2003).

Modelul bazat pe mediile "celulelor" este, la rândul său, împărţit de Kirk (2003) în două tipuri:

- *Unweighted means*, ce vizează simpla medie aritmetică a condiţiilor experimentului.
- *Weighted means*, ce vizează media condiţiilor în funcţie de mărimea eşantionului - *n(j, k)*.

Pentru o înţelegere aprofundată a conceptului de analiză a mediilor *unweighted*, se va face referire la articolul lui Mistri şi al colaboratorilor săi (2020).

## Design Factorial Aleatoriu în Block-uri

Consensul metodologic privind notarea acestui tip de design este RBF-*pq*, unde:

- RBF = “Randomized Block Factorial”
- $p$ = Numărul de niveluri ale intervenţiei A.
- $q$ = Numărul de niveluri ale intervenţiei B.

Acest tip de design este obţinut prin combinarea nivelurilor unui design RB-$p$ cu nivelurile unui design RB-$q$, astfel încât fiecare nivel al design-ului RB-$p$ apare o singură dată asociat cu fiecare nivel al design-ului RB-$q$ şi reciproc. Design-ul utilizează tehnica stratificării (i.e., dispunerea în block-uri, *blocking*), împreună cu un design RB-$p$, pentru a izola varianţa variabilelor-zgomot. Simultan, pot fi evaluate 2 sau mai multe tratamente şi interacţiuni asociate (Kirk, 2003).

Pentru o înţelegere aprofundată a design-ului factorial aleatoriu în block-uri, se va face referire atât la schemele generale propuse de Kirk (2003, pp. 20-21), cât şi la un articol din literatura de specialitate (Lim, Bong, Kang, & Dennen, 2019), după cum urmează.

| | Treatment Combination | Treatment Combination | Treatment Combination | Treatment Combination | Treatment Combination | Treatment Combination | |
|---|---|---|---|---|---|---|---|
| $\text{Block}_1$ | $a_1b_1$ | $a_1b_2$ | $a_2b_1$ | $a_2b_2$ | $a_3b_1$ | $a_3b_2$ | $\overline{Y}_{1..}$ |
| $\text{Block}_2$ | $a_1b_1$ | $a_1b_2$ | $a_2b_1$ | $a_2b_2$ | $a_3b_1$ | $a_3b_2$ | $\overline{Y}_{2..}$ |
| $\text{Block}_3$ | $a_1b_1$ | $a_1b_2$ | $a_2b_1$ | $a_2b_2$ | $a_3b_1$ | $a_3b_2$ | $\overline{Y}_{3..}$ |
| $\text{Block}_4$ | $a_1b_1$ | $a_1b_2$ | $a_2b_1$ | $a_2b_2$ | $a_3b_1$ | $a_3b_2$ | $\overline{Y}_{4..}$ |
| $\text{Block}_5$ | $a_1b_1$ | $a_1b_2$ | $a_2b_1$ | $a_2b_2$ | $a_3b_1$ | $a_3b_2$ | $\overline{Y}_{5..}$ |
| | $\overline{Y}_{.11}$ | $\overline{Y}_{.12}$ | $\overline{Y}_{.21}$ | $\overline{Y}_{.22}$ | $\overline{Y}_{.31}$ | $\overline{Y}_{.32}$ | |

**Figure 1.10** Layout for a two-treatment randomized block factorial design (RBF-32 design). Each block contains six girls who are overweight by about the same amount. The girls in a block are randomly assigned to the six treatment combinations.

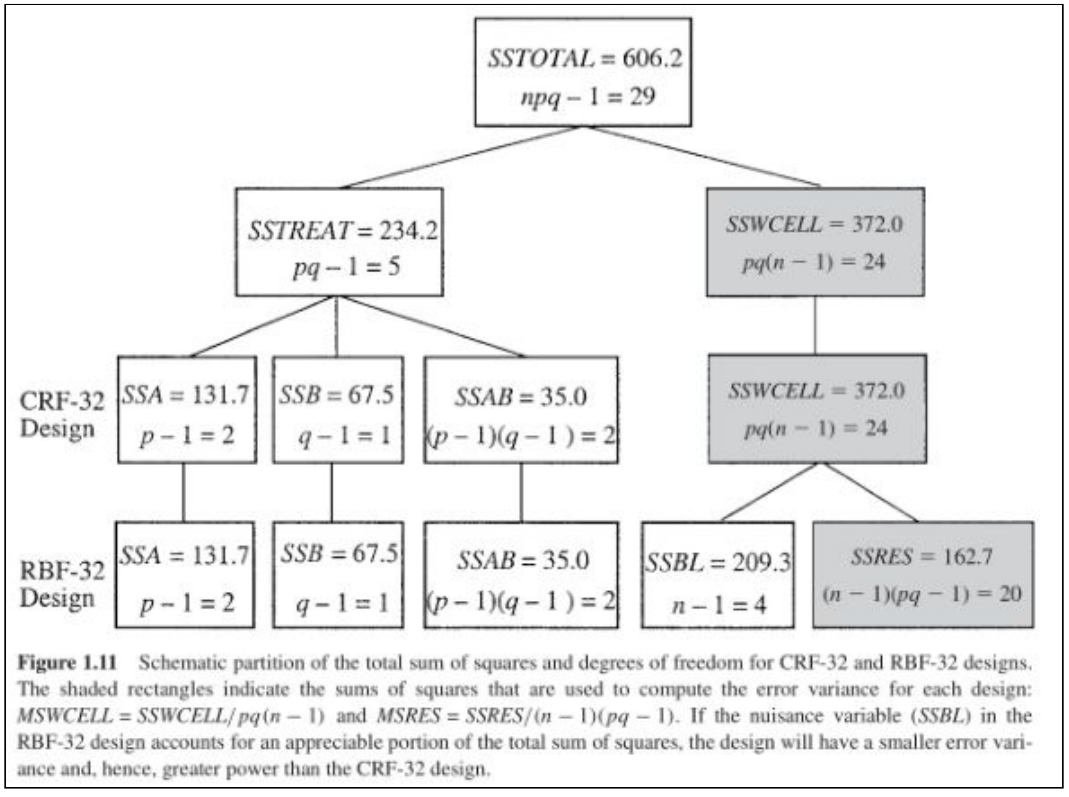

SSTOTAL = 606.2
npq − 1 = 29
SSTREAT = 234.2
pq − 1 = 5
SSWCELL = 372.0
pq(n − 1) = 24
CRF-32 Design
SSA = 131.7
p − 1 = 2
SSB = 67.5
q − 1 = 1
SSAB = 35.0
(p − 1)(q − 1) = 2
SSWCELL = 372.0
pq(n − 1) = 24
RBF-32 Design
SSA = 131.7
p − 1 = 2
SSB = 67.5
q − 1 = 1
SSAB = 35.0
(p − 1)(q − 1) = 2
SSBL = 209.3
n − 1 = 4
SSRES = 162.7
(n − 1)(pq − 1) = 20

**Figure 1.11** Schematic partition of the total sum of squares and degrees of freedom for CRF-32 and RBF-32 designs. The shaded rectangles indicate the sums of squares that are used to compute the error variance for each design: $MSWCELL = SSWCELL/pq(n-1)$ and $MSRES = SSRES/(n-1)(pq-1)$. If the nuisance variable ($SSBL$) in the RBF-32 design accounts for an appreciable portion of the total sum of squares, the design will have a smaller error variance and, hence, greater power than the CRF-32 design.

# Design-uri Factoriale cu Confundate/ Covariate

## Fundamente Teoretice

În cazul în care dimensiunea unui block este prea mare, o formă alternativă de adresare a subiectului de cercetare este un design factorial confundat cu două tipuri de tratament. Acest din urmă tip de design plasează covariata în interacţiunea dintre anumite grupuri de block-uri. Drept rezultat, testul pe interacţiune este mai puţin puternic decât testele pe tratamentele A şi B. Aceste tipuri de design-uri sunt construite fie dintr-un design aleatoriu în block-uri, fie dintr-un design de tip pătrat latin (Kirk, 2003). Consensul metodologic de notare a acestor design-uri este RBCF-$p^k$ şi LSCF-$p^k$, unde:

- RB = "Randomized Block".
- LS = "Latin Square".
- C = Interacţiunea este complet confundată cu grupurile de block-uri.
- F = Design factorial.
- $p^k$ = Design-ul are $k$ tratamente/ intervenţii, fiecare pe $p$ niveluri.

Pentru detalierea argumentelor anterior stipulate, se va face referire la schemele generale propuse de Kirk (2003, pp. 24-25), după cum urmează.

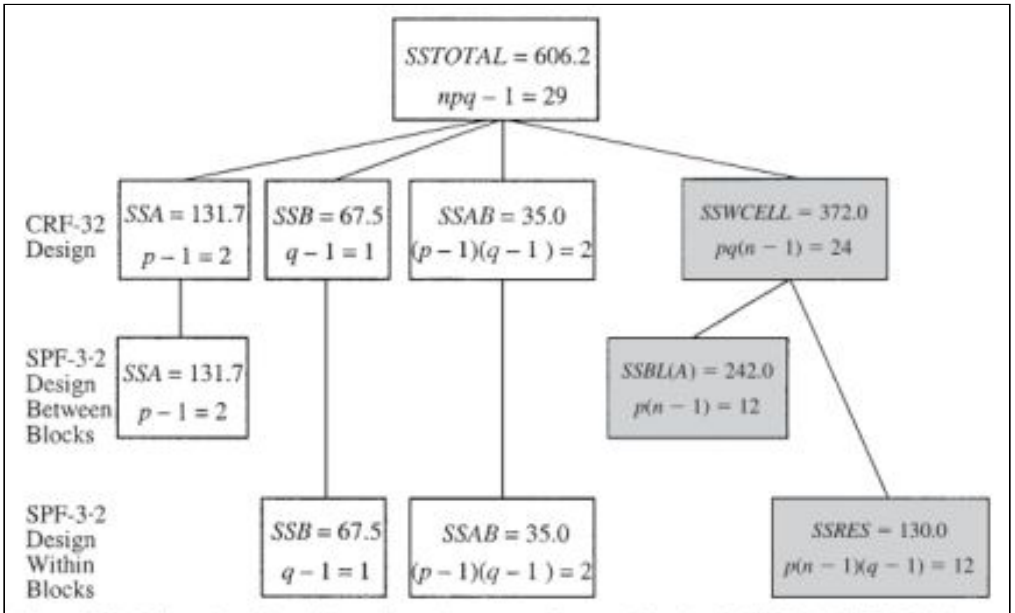

SSTOTAL = 606.2
npq − 1 = 29
CRF-32 Design
SSA = 131.7
p − 1 = 2
SSB = 67.5
q − 1 = 1
SSAB = 35.0
(p − 1)(q − 1) = 2
SSWCELL = 372.0
pq(n − 1) = 24
SPF-3·2 Design Between Blocks
SSA = 131.7
p − 1 = 2
SSBL(A) = 242.0
p(n − 1) = 12
SPF-3·2 Design Within Blocks
SSB = 67.5
q − 1 = 1
SSAB = 35.0
(p − 1)(q − 1) = 2
SSRES = 130.0
p(n − 1)(q − 1) = 12

**Figure 1.13** Schematic partition of the total sum of squares and degrees of freedom for CRF-32 and SPF-3·2 designs. The shaded rectangles indicate the sums of squares that are used to compute the error variance for each design. The SPF-3·2 design has two error variances: $MSBL(A) = SSBL(A)/p(n-1)$ is used to test treatment $A$; $MSRES = SSRES/p(n-1)(q-1)$ is used to test treatment $B$ and the $A \times B$ interaction. The within-blocks error variance, $MSRES$, is usually much smaller than the between-blocks error variance, $MSBL(A)$. As a result, tests of treatment $B$ and the $A \times B$ interaction are more powerful than the test of treatment $A$.

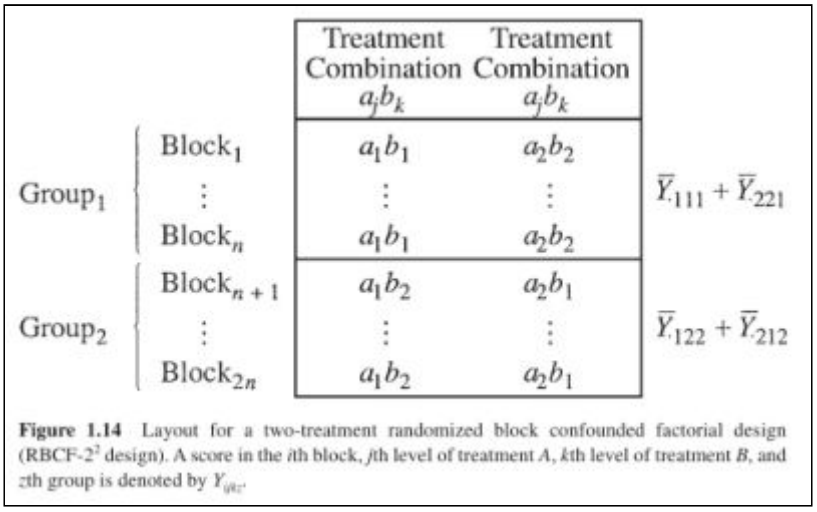



**Figure 1.14** Layout for a two-treatment randomized block confounded factorial design (RBCF-$2^2$ design). A score in the $i$th block, $j$th level of treatment $A$, $k$th level of treatment $B$, and $z$th group is denoted by $Y_{ijkz}$.

## Design Factorial de Tip Split-Plot

Acest tip de design provine din studiile agriculturale, în care tratamentul A este aplicat la parcele relativ mari - la totalitatea parcelelor. Totalitatea parcelelor este, apoi, subdivizată (*split*), urmând ca nivelurile tratamentului B să fie aplicate la subdiviziunile parcelelor respective. În acest sens, tratamentul A este considerat unul *between-blocks*, în timp ce tratamentul B este considerat unul *within-blocks* (Kirk, 2003).

Schema generală (Kirk, 2003, p. 22), acompaniată de recomandarea unui articol elocvent (Purdum, Eusebio, & Hanford, 2020), este ilustrată în continuare.

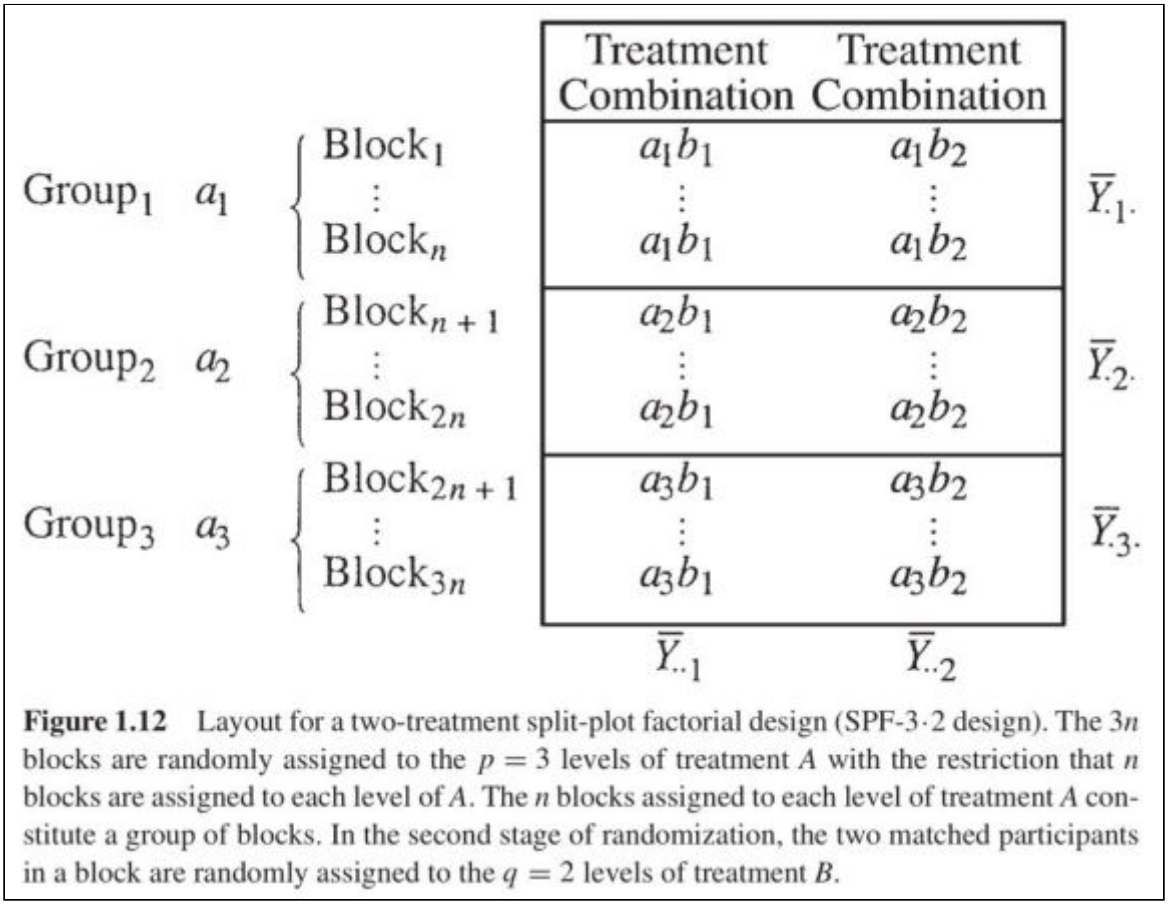



**Figure 1.12** Layout for a two-treatment split-plot factorial design (SPF-3·2 design). The $3n$ blocks are randomly assigned to the $p = 3$ levels of treatment $A$ with the restriction that $n$ blocks are assigned to each level of $A$. The $n$ blocks assigned to each level of treatment $A$ constitute a group of blocks. In the second stage of randomization, the two matched participants in a block are randomly assigned to the $q = 2$ levels of treatment $B$.

### Design-uri Factoriale Fracţionare

Aceste design-uri reduc numărul de combinări posibile dintre tratamentele ce pot fi incluse într-un experiment de tip multi-tratament la fracţii ale numărului total de combinări. În teorie, sunt similare cu design-urile factoriale confundate. Cu toate acestea, limitarea acestei reduceri a numărului de combinări posibile ale tratamentelor constă în ambiguitatea de interpretare a rezultatelor unui astfel de experiment (i.e., deoarece 2 sau mai multe denumiri pot fi atribuite fiecărei sume de pătrate) (Kirk, 2003).

Pentru exemplificare, schema generală propusă de Kirk (2003, p. 26), urmată de o recomandare de articol (El-Helaly, Habib, & Abd El-Rahman, 2018), va fi ilustrată.

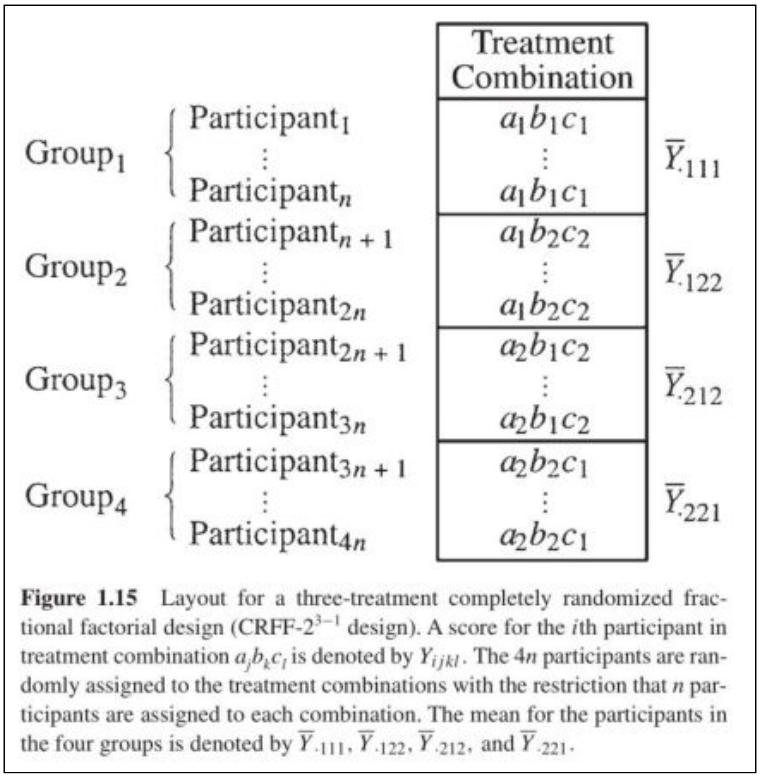



**Figure 1.15** Layout for a three-treatment completely randomized fractional factorial design (CRFF-$2^{3-1}$ design). A score for the $i$th participant in treatment combination $a_j b_k c_l$ is denoted by $Y_{ijkl}$. The $4n$ participants are randomly assigned to the treatment combinations with the restriction that $n$ participants are assigned to each combination. The mean for the participants in the four groups is denoted by $\overline{Y}_{\cdot 111}$, $\overline{Y}_{\cdot 122}$, $\overline{Y}_{\cdot 212}$, and $\overline{Y}_{\cdot 221}$.

## Design-uri Ierarhice

### Design-uri Ierarhice cu Unul/ Două Tratamente Grupate

Aceste tipuri de design-uri sunt construite fie din design-uri complet aleatorii, fie din design-uri aleatorii în block-uri. Un design ierarhic cu două tratamente este construit dintr-un design CR-*p* şi un design CR-*q* şi este marcat prin abrevierea CRH-*pq(A)* (Kirk, 2003), unde:

- CRH = "Completely Randomized Hierarchical".
- *pq(A)* = Design-ul are *p* niveluri de tratament A şi *q* niveluri de tratament B care este grupat (*nested*) în tratamentul A.

Pentru exemplificare, se va recurge la schemele generale propuse de Kirk (2003, pp. 27-28), urmate de recomandarea unui articol elocvent (Martindale, 1978), în acest sens.

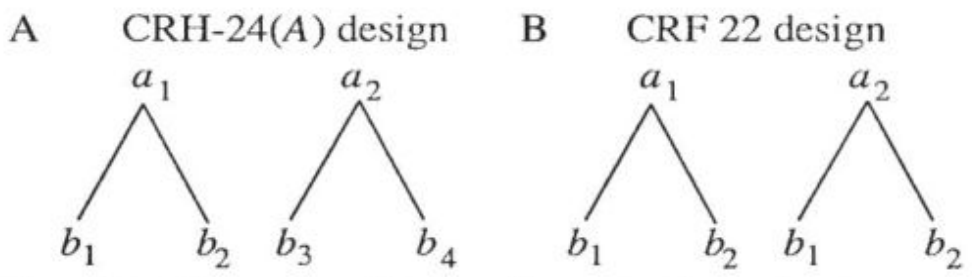



**Figure 1.16** Comparison of designs with nested and crossed treatments. In panel A, treatment $B(A)$ is nested in treatment $A$ because $b_1$ and $b_2$ appear only with $a_1$ while $b_3$ and $b_4$ appear only with $a_2$. In panel B, treatments $A$ and $B$ are crossed because each level of treatment $B$ appears once and only once with each level of treatment $A$ and vice versa.

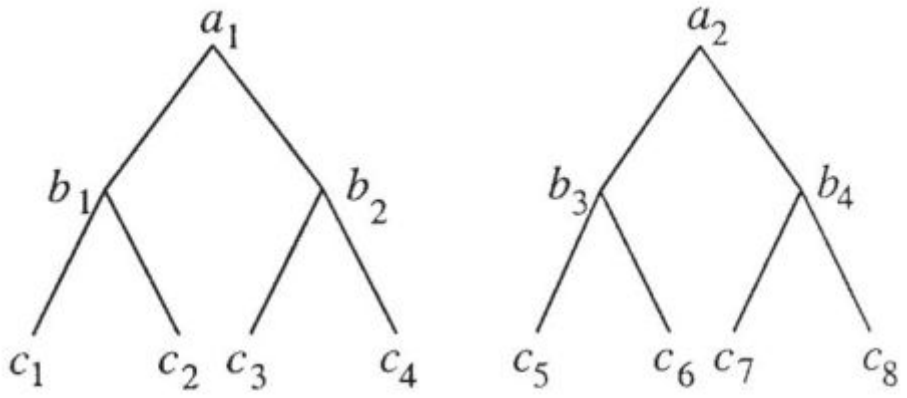



**Figure 1.17** Diagram of a three-treatment completely randomized hierarchical design (CRH-24(*A*)8(*AB*) design). The four schools, $b_1, \ldots, b_4$, are nested in the two approaches to introducing long division, treatment $A$. The eight teachers, $c_1, \ldots, c_8$, are nested in the schools and teaching approaches. Students are randomly assigned to the $pq_{(j)}r_{(jk)} = (2)(2)(2) = 8$ treatment combinations with the restriction that $n$ students are assigned to each combination.

## Design Ierarhic cu Tratamente Încrucișate și Grupate

Acest tip de design este similar cu cel discutat anterior. În acest sens, se va trece direct la exemplificare, cu ajutorul schemei generale propuse de Kirk (2003, p. 29) și cu ajutorul unui articol (Carter, Holmes III, Fletcher, & McCollough, 2019).

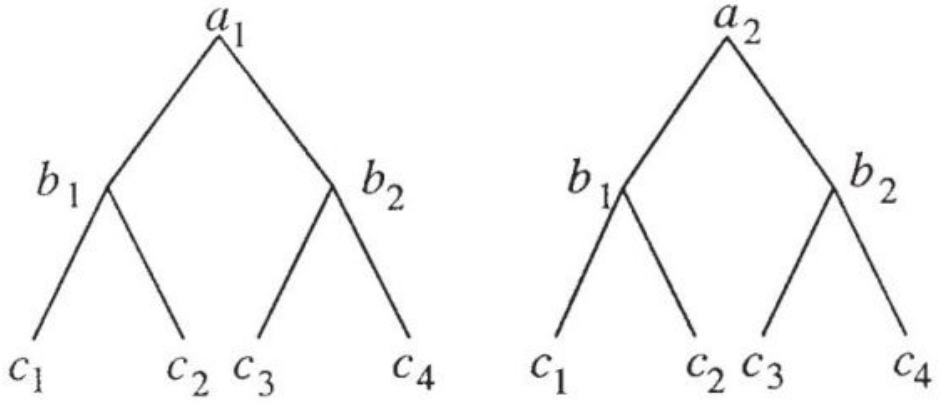



**Figure 1.18** Diagram of a three-treatment completely randomized partial hierarchical design (CRPH-*pqr*(*B*) design). The letter P in the designation stands for "partial" and indicates that not all of the treatments are nested. In this example, treatments *A* and *B* are crossed; treatment *C*(*B*) is nested in treatment *B* because $c_1$ and $c_2$ appear only with $b_1$ while $c_3$ and $c_4$ appear only with $b_2$. Treatment *C*(*B*) is crossed with treatment *A* because each level of treatment *C*(*B*) appears once and only once with each level of treatment *A* and vice versa.

# Design-uri Experimentale cu O Covariată

## Analiza de Covarianţă (*ANCOVA*)

Analiza de covarianţă reprezintă o modalitate alternativă de reducere a erorilor de varianţă şi de minimizare a efectelor variabilelor confundate. Aceasta combină analiza de regresie şi analiza de varianţă. Totodată, controlul experimental poate lua forme multiple (e.g., alocare aleatorie a participanţilor la nivelurile de tratament, stratificarea participanţilor în block-uri omogene, rafinarea tehnicilor de măsurare a variabilei dependente). Deseori, analiza de covarianţă este utilizată într-unul din trei situaţii de cercetare (Kirk, 2003):

- Utilizarea grupurilor intacte cu medii inegale ale variabilelor experimentale, în cercetarea educaţională şi industrială.
- Atunci când devine aparent că, în ciuda faptului că a fost utilizată alocarea aleatorie, participanţii au fost neechivalenţi pe o variabilă relevantă, la începutul experimentului.
- Au loc diferenţe în variabilele-zgomot relevante pentru experiment.

Este important de menţionat faptul că controlul statistic şi controlul experimental nu sunt accepţii mutual exclusive de reducere a erorilor de varianţă şi de minimizare a efectelor variabilelor-zgomot (Kirk, 2003).

Ca o încheiere, Kirk (2003) afirmă că design-urile experimentale diferă printr-o serie de modalităţi posibile, precum:

- Procedurile de selecţie şi de alocare aleatorie.
- Numărul de tratamente/ intervenţii.
- Utilizarea eşantioanelor independente sau dependente cu stratificare.
- Utilizarea de tratamente încrucişate şi grupate.

- Prezenţa confundării.
- Utilizarea covariatelor.

În scopul exemplificării procedeului analizei de covarianţă, se va face apel la schema generală a acesteia (Kirk, 2003, pp. 29-30) şi la un articol din literatura de specialitate (Szczesniak et al., 2020).

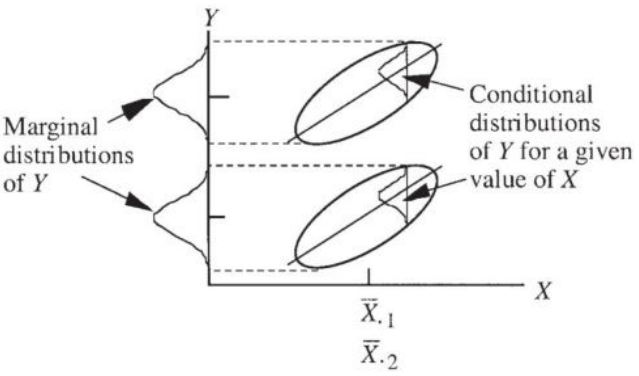



**Figure 1.19** Scatter plots showing the relationship between the dependent variable, *Y*, and concomitant variable, *X*, for the two treatment levels. The size of the error variance in ANOVA is determined by the dispersion of the marginal distributions. The size of the error variance in ANCOVA is determined by the dispersion of the conditional distributions. The higher the correlation between *X* and *Y* is, the greater the reduction in the error variance due to using analysis of covariance.

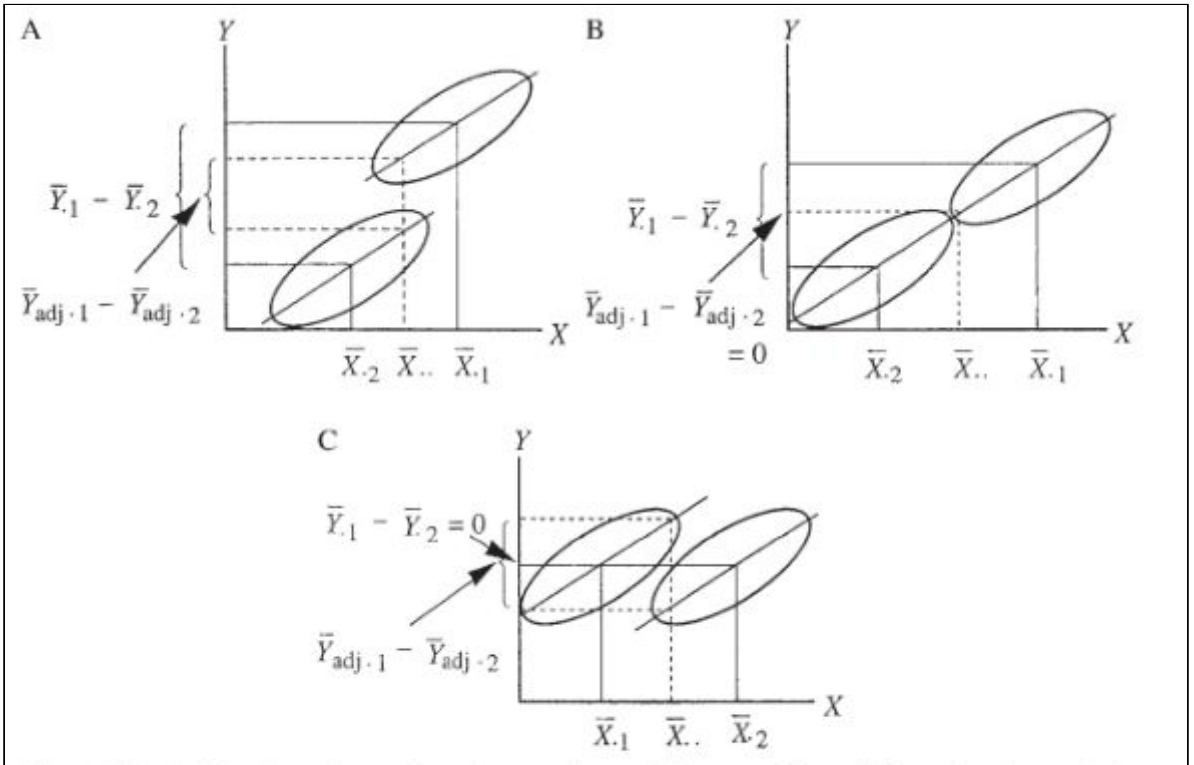



**Figure 1.20** Analysis of covariance adjusts the concomitant-variable means, $\overline{X}_{\cdot 1}$ and $\overline{X}_{\cdot 2}$, so that they equal the concomitant-variable grand mean, $\overline{X}_{\cdot\cdot}$. When the concomitant-variable means differ, the absolute difference between adjusted means for the dependent variable, $|\overline{Y}_{\text{adj}\cdot 1} - \overline{Y}_{\text{adj}\cdot 2}|$, can be less than that between unadjusted means, $|\overline{Y}_{\cdot 1} - \overline{Y}_{\cdot 2}|$, as in panels A and B, or larger, as in panel C.

## Sumar - Design-uri de Cercetare Experimentale

- Analiză robustă a naturii/ realităţii.
- Decizii informate.
- Selectare/ Alocare aleatorie + Controlul variabilei dependente + Controlul covariatelor.
- Design complet aleatoriu.
- Design aleatoriu stratificat (în block-uri, RB-uri).
- Design de tip pătrat latin (LS-uri).
- Factori = Variabile independente majore.
- *Omega Squared.*
- *Blocking* = Stratificare (i.e., izolează varianţa variabilelor-zgomot).
- Design-urile factoriale cu confundate/ covariate reprezintă alternative la design-urile în block-uri, atunci când block-urile sunt prea mari.
- Design-urile factoriale cu confundate/ covariate sunt fie RB-uri, fie LS-uri.
- Design-urile ierarhice sunt fie design-uri complet aleatorii, fie RB-uri.
- În design-urile ierarhice, tratamentele secundare sunt grupate (*nested*) în tratamentele iniţiale.
- *ANCOVA* = Regresie + *ANOVA*.
- *ANCOVA* asigură controlul variabilelor confundate.

## Bibliografie Selectivă - Design-uri de Cercetare Experimentale


Adam, R. S. Y., & Hadi, Y. (2018). The effect of writing format on student reading comprehension on paper media using completely randomized factorial design. *Journal of Engineering and Management in Industrial System*, *6*(2), 95-104. https://doi.org/10.21776/ub.jemis.2018.006.02.4

Carter, R. E., Holmes III, D. R., Fletcher, J. G., & McCollough, C. H. (2019). Evaluation of Pseudoreader Study Designs to Estimate Observer Performance Results as an Alternative to Fully Crossed, Multireader, Multicase Studies. *Academic Radiology*, *27*(2), 244-252. https://doi.org/10.1016/j.acra.2019.03.014

Chen, W. L., Ye, Q., Zhang, S. C., Xia, Y., Yang, X., Yuan, T. F., ... & Li, J. A. (2019). Aphasia rehabilitation based on mirror neuron theory: a randomized-block-design study of neuropsychology and functional magnetic resonance imaging. *Neural Regeneration Research*, *14*(6), 1004-1012. https://doi.org/10.4103/1673-5374.250580

da Silva, I. C. B., Fernandes, L. A., Colen, F., & Sampaio, R. A. (2017). Growth and production of common bean fertilized with biochar. *Ciência Rural*, *47*(11), 1-8. https://doi.org/10.1590/0103-8478cr20170220

El-Helaly, S. N., Habib, B. A., & Abd El-Rahman, M. K. (2018). Resolution V fractional factorial design for screening of factors affecting weakly basic drugs liposomal systems. *European Journal of Pharmaceutical Sciences*, *119*, 249-258. https://doi.org/10.1016/j.ejps.2018.04.028

Lim, T., Bong, J., Kang, J. H., & Dennen, V. (2019). The Effects of Information Volume and Distribution on Cognitive Load and Recall: Implications for the Design of Mobile Marker-less Augmented Reality. *Educational Technology International*, *20*(2), 137-168. Available at http://210.101.116.28/kiss10/preview.asp

Martindale, C. (1978). The therapist-as-fixed-effect fallacy in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *46*(6), 1526–1530. https://doi.org/10.1037/0022-006X.46.6.1526

Mistri, M., Munari, C., Pagnoni, A., Chenet, T., Pasti, L., & Cavazzini, A. (2020). Accumulation of trace metals in crayfish tissues: is Procambarus clarkii a vector of pollutants in Po Delta inland waters?. *The European Zoological Journal*, *87*(1), 46-57. https://doi.org/10.1080/24750263.2020.1717653

Pine, P. S., Munro, S. A., Parsons, J. R., McDaniel, J., Lucas, A. B., Lozach, J., ... & Salit, M. (2016). Evaluation of the External RNA Controls Consortium (ERCC) reference material using a modified Latin square design. *BMC Biotechnology*, *16*(1), 1-15. https://doi.org/10.1186/s12896-016-0281-x

Purdum, S., Eusebio, P., & Hanford, K. (2020). The effects of two genetic lines on spatial distribution and use and preference of perch and nest area in an aviary system. *Poultry Science*, *99*, 3328-3333. https://doi.org/10.1016/j.psj.2020.03.056

Szczesniak, D., Ciulkowicz, M., Maciaszek, J., Misiak, B., Luc, D., Wieczorek, T., ... & Rymaszewska, J. (2020). Psychopathological responses and face mask restrictions during the COVID-19 outbreak: Results from a nationwide survey. *Brain, Behavior, and Immunity*, *87*, 161-162. https://doi.org/10.1016%2Fj.bbi.2020.05.027

## Resurse Online - Design-uri de Cercetare Experimentale

1. Experimental Design
2. A Quick Guide to Experimental Design | 4 Steps & Examples
3. Experimental Design | Research Methods Knowledge Base
4. Experimental Design
5. https://alexandercoppock.com/menu_statistical_tools.html

# Design-uri de Cercetare Cvasi-Experimentale

Înainte de a începe acest capitol, vor fi listate notaţiile de care va fi nevoie pentru înţelegerea schemelor generale ale modelelor de design-uri de cercetare cvasi-experimentale, după cum urmează:

- *M* = Materiale/ Instrumente utilizate pentru măsurare.
- *O* = Observaţii/ Măsurători.
- *R* = Alocare aleatorie (*Randomization*).
- *X* = Variabilă experimentală/ Intervenţie/ Tratament.

## Experiment pe Serii de Timp (*Time-Series Experiment*)

Experimentul pe serii de timp constă în prezenţa unui proces periodic de măsurare a unui grup sau a unui individ. Acest proces constă într-o introducere a unei schimbări experimentale în seria de timp, ale cărei rezultate sunt indicate prin discontinuitatea măsurătorilor din seria respectivă (Campbell & Stanley, 2015).

Schema generală a modelului, ilustrată în continuare, alături de un articol (Rosmalen, Wenting, Roest, de Jonge, & Bos, 2012) vor permite înţelegerea aprofundată a acestui design de cercetare.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $O_1$ | $O_2$ | $O_3$ | $O_4$ | $X$ | $O_5$ | $O_6$ | $O_7$ | $O_8$ |

## Design pe Eşantioane Temporale Echivalente (*Equivalent Time-Sample Design*)

Acest design reprezintă o formă de experimentare recurentă pe un singur grup. În acest sens, se selectează două eşantioane echivalente din grupul respectiv. Într-unul dintre eşantioane, variabila experimentală este prezentă, în timp ce în celălalt eşantion nu este (Campbell & Stanley, 2015).

Schema generală a modelului, ilustrată în continuare, alături de un articol (Hayward et al., 2017) vor permite înţelegerea aprofundată a acestui design de cercetare.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $X_1$ | $O$ | | $X_0$ | $O$ | | $X_1$ | $O$ | | $X_0$ | $O$ |

## Design cu Materiale/ Instrumente Echivalente (*Equivalent Materials Design*)

Argumentul echivalenţei eşantioanelor selectate este bazat pe materialele sau instrumentele cu ajutorul cărora sunt comparate variabilele experimentale (Campbell & Stanley, 2015).

Schema generală a modelului, ilustrată în continuare, alături de un articol (Jetter, 1977, pp. 21-28) vor permite înţelegerea aprofundată a acestui design de cercetare.

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $M_a$ | $X_1$ | $O$ | | $M_b$ | $X_0$ | $O$ | | $M_c$ | $X_1$ | $O$ | | $M_d$ | $X_0$ | $O$ |

## Design Neechivalent cu Grup de Control (*Nonequivalent Control Group Design*)

În cadrul acestui tip de design, grupurile constituie colective formate natural (e.g., clase de elevi), pe cât de similare posibil, dar nu într-atât de similare încât să poată lipsi pretestarea. Totodată, alocarea variabilei experimentale într-unul din grupuri este aleatorie (Campbell & Stanley, 2015).

Schema generală a modelului, ilustrată în continuare, alături de un articol (Kenny, 1975) vor permite înțelegerea aprofundată a acestui design de cercetare.

| | | |
|---|---|---|
| *O* | *X* | *O* |
| *O* | | *O* |

## Design-uri Contrabalansate (*Counterbalanced Designs*)

Aceste design-uri reprezintă design-uri prin care controlul experimental este obținut cu ajutorul alocării tuturor respondenților sau a condițiilor în toate formele de tratament sau intervenție (Campbell & Stanley, 2015).

Schema generală a modelului, ilustrată în continuare, alături de un articol (Pollatek & Well, 1995) vor permite înțelegerea aprofundată a acestui design de cercetare.

| | **Timp 1** | **Timp 2** | **Timp 3** | **Timp 4** |
|---|---|---|---|---|
| **Grup A** | $X_1O$ | $X_2O$ | $X_3O$ | $X_4O$ |
| **Grup B** | $X_2O$ | $X_4O$ | $X_1O$ | $X_3O$ |
| **Grup C** | $X_3O$ | $X_1O$ | $X_4O$ | $X_2O$ |
| **Grup D** | $X_4O$ | $X_3O$ | $X_2O$ | $X_1O$ |

## Design cu Eşantion Separat Pretest-Posttest (*Separate Sample Pretest-Posttest Design*)

Acest tip de design este utilizat în cazul în care, pentru populaţii mari (e.g., oraşe, fabrici, etc.), un cercetător poate încerca să controleze experimental condiţiile de "*când*" şi "*cui* se adresează" măsurătorile, prin proceduri de alocare aleatorie. În ciuda acestui fapt, este dificilă separarea de subgrupuri, în sensul administrării unor tratamente experimentale diferite (Campbell & Stanley, 2015).

Schema generală a modelului, ilustrată în continuare, alături de un articol (Markert, O'Neill, & Bhatia, 2003) vor permite înțelegerea aprofundată a acestui design de cercetare.

| | | | |
|---|---|---|---|
| *R* | *O* | *(X)* | |
| *R* | | *X* | *O* |

## Design cu Eşantion Separat Pretest-Posttest şi Grup de Control (*Separate Sample Pretest-Posttest Control Group Design*)

Acest tip de design este similar cu design-ul cu eşantion separat pretest-posttest, discutat anterior. În cazul în care grupurile din care variabila experimentală poate fi păstrată sunt comparabile (i.e., neechivalente), atunci poate fi adăugat un grup de control (Campbell & Stanley, 2015).

În continuare, va fi ilustrată schema generală a modelului.

| | | | |
|---|---|---|---|
| *R* | *O* | *(X)* | |
| *R* | | *X* | *O* |
| *R* | *O* | | |
| *R* | | | *O* |

## Design pe Serii Multiple de Timp (*Multiple Time-Series Design*)

Acest tip de design este util în studierea schimbărilor majore, în timp. Astfel, cercetătorul caută o instituţie similară care nu a implementat variabila experimentală X, drept "grup de control". Adiţional, este ideal ca X să fie alocat aleatoriu (Campbell & Stanley, 2015).

Schema generală a modelului, ilustrată în continuare, alături de un articol (Chung, Wan, Kuo, Lin, & Liu, 2018, pp. 1-8) vor permite înţelegerea aprofundată a acestui design de cercetare.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *O* | *O* | *O* | *O* | *X* | *O* | *O* | *O* | *O* |
| *O* | *O* | *O* | *O* | | *O* | *O* | *O* | *O* |

## Design-ul Ciclului Instituţional Recurent (*Recurrent Institutional Cycle Design*)

Acest tip de design este cunoscut şi drept un design "*patched-up*" (Campbell & Stanley, 2015, p. 57), prin prisma faptului că un cercetător porneşte la drum cu un design neadecvat şi adaugă, pe parcurs, elemente specifice (e.g., alte măsurători), cu scopul de a controla erori recurente de validitate.

În cele ce urmează, vor fi delimitate două tipuri de scheme generale (i.e., un model ideal şi un model real, de tip studiu de caz), alături de un articol (Di Noia, Orr, & Byrd-Bredbenner, 2014) din literatura de specialitate, pentru aprofundare.

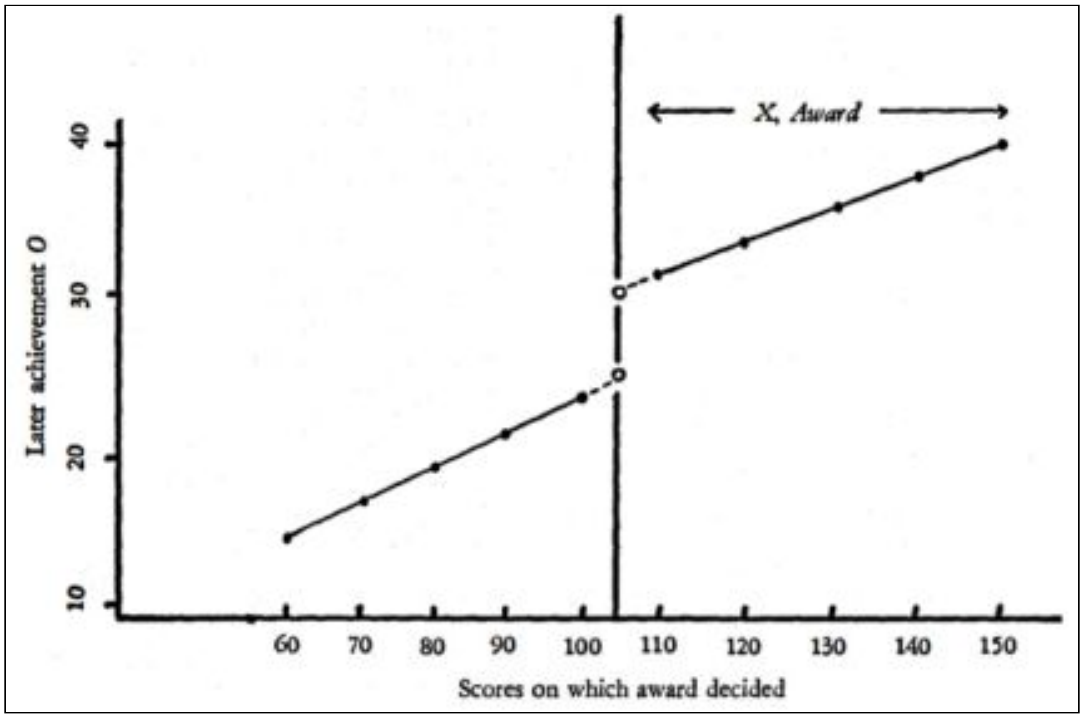
| Model Ideal | | | | |
|---|---|---|---|---|
| **Eşantion A** | *X* | $O_1$ | | |
| **Eşantion B** | | $O_2$ | *X* | $O_3$ |

| Model Real (Studiu de Caz - Școală) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Clasa A** | $X$ | $O_1$ | | | | | | |
| **Clasa B1** | | | $R$ | $O_2$ | $X$ | $O_3$ | | |
| **Clasa B2** | | | $R$ | | $X$ | $O_4$ | | |
| **Clasa C** | | | | | | | $O_5$ | $X$ |

## Analiza Regresie-Discontinuitate (*Regression-Discontinuity Analysis*)

Analiza regresie-discontinuitate este utilizată, majoritar, în scopuri educaţionale. Design-ul respectiv este dezvoltat în situaţia în care au fost deja utilizate analize *ex post facto* (i.e., analize în care investigaţia începe abia după ce un anumit eveniment a avut loc, fără implicare din partea cercetătorului) (Campbell & Stanley, 2015; Salkind, 2010b).

Schema generală a modelului, propusă de Campbell şi Stanley (2015, p. 62), alături de un articol (Lee, 2001) vor permite înţelegerea aprofundată a acestui design de cercetare.



## Sumar - Design-uri de Cercetare Cvasi-Experimentale

- Experimentele pe serii de timp presupun măsurători longitudinale, în cadrul cărora este prezent un singur tip de intervenţie.
- Design-ul pe eşantioane temporale echivalente presupune introducerea repetată a variabilei experimentale. De asemenea, unul dintre eşantioane simulează grupul de control.
- Design-ul cu instrumente echivalente presupune compararea intervenţiilor în funcţie de instrumentele utilizate în măsurare.
- Design-urile neechivalente cu grup de control presupun alocarea aleatorie a variabilei experimentale. În acest sens, grupul experimental primeşte intervenţia X, în timp ce grupul de control nu primeşte intervenţia X. Este important de menţionat faptul că aceste grupuri sunt formate natural.
- Design-urile contrabalansate sunt utilizate pentru compararea diverselor intervenţii, în cazul grupurilor multiple, în momente de timp multiple. Totodată, alocările se fac prin rotaţie.
- Un design cu eşantion separat pretest-posttest se poate dovedi de interes pentru controlul experimental asupra unor populaţii mari. În cadrul acestui tip de design, alocarea participanţilor în grupurile de interes este aleatorie (*R*). Intervenţia este opţională în primul grup, dacă se regăseşte în cel de-al doilea.
- Principiile design-ului cu eşantion separat pretest-posttest pot fi extrapolate la acelaşi tip de design cu grup de control (i.e., grup în care participanţii nu primesc intervenţie) adiţional.
- În cadrul design-ului pe serii multiple de timp, grupurile diferă prin prezenţa sau absenţa lui *X*. Modelul design-ului cuprinde două grupuri şi măsurători în momente similare de timp. De asemenea, intervenţia (*X*) ar trebui să fie alocată aleatoriu.
- În cadrul unui design de tip ciclu instituţional recurent, se porneşte la drum cu un design neadecvat, adăugându-se, pe parcurs, elemente specifice, cu ajutorul cărora cercetătorul compensează validitatea scăzută a studiului.
- Analiza regresie-discontinuitate vizează punctul de discontinuitate în varianţa variabilei dependente (VD) (e.g., scorurile unui test), respectiv amplificarea sau diminuarea efectului variabilei independente (VI) (e.g., timp).
- Explicaţia simplă a analizei regresie-discontinuitate: “Intervenţia a avut vreun efect, în timp?”.
- Explicaţia mai complexă a analizei regresie-discontinuitate: “Dacă VD rămâne pe loc, în timp ce VI creşte/ scade, ce se va întâmpla cu VD, la un anumit moment? Va creşte sau va scădea, concomitent cu VI?”.

## Bibliografie Selectivă - Design-uri de Cercetare Cvasi-Experimentale


Chung, F. F., Wan, G. H., Kuo, S. C., Lin, K. C., & Liu, H. E. (2018). Mother–infant interaction quality and sense of parenting competence at six months postpartum for first-time mothers in Taiwan: a multiple time series design. *BMC Pregnancy and Childbirth*, *18*(1), 1-13. https://doi.org/10.1186/s12884-018-1979-7

Di Noia, J., Orr, L., & Byrd-Bredbenner, C. (2014). Residential summer camp intervention improves camp food environment. *American Journal of Health Behavior*, *38*(4), 631-640. https://doi.org/10.5993/AJHB.38.4.17

Hayward, K. L., Martin, J. H., Cottrell, W. N., Karmakar, A., Horsfall, L. U., Patel, P. J., ... & Valery, P. C. (2017). Patient-oriented education and medication management intervention

for people with decompensated cirrhosis: study protocol for a randomized controlled trial. *Trials*, *18*(1), 1-13. https://doi.org/10.1186/s13063-017-2075-4

Jetter, J. T. (1977). A systematic approach to aural-visual identification instruction in music for young children. *Journal of Research in Music Education*, *25*(1), 21-31. https://doi.org/10.2307%2F3344842

Kenny, D. A. (1975). A quasi-experimental approach to assessing treatment effects in the nonequivalent control group design. *Psychological Bulletin*, *82*(3), 345–362. https://doi.org/10.1037/0033-2909.82.3.345

Lee, D. S. (2001). *The Electoral Advantage to Incumbency and Voters' Valuation of Politicians' Experience: A Regression Discontinuity Analysis of Elections to the US House* (pp. 1-43). (No. w8441). National Bureau of Economic Research. https://doi.org/10.3386/w8441

Markert, R. J., O'Neill, S. C., & Bhatia, S. C. (2003). Using a quasi-experimental research design to assess knowledge in continuing medical education programs. *Journal of Continuing Education in the Health Professions*, *23*(3), 157-161. https://doi.org/10.1002/chp.1340230306

Pollatsek, A., & Well, A. D. (1995). On the use of counterbalanced designs in cognitive research: A suggestion for a better and more powerful analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(3), 785-794. https://doi.org/10.1037/0278-7393.21.3.785

Rosmalen, J. G., Wenting, A. M., Roest, A. M., de Jonge, P., & Bos, E. H. (2012). Revealing causal heterogeneity using time series analysis of ambulatory assessments: application to the association between depression and physical activity after myocardial infarction. *Psychosomatic Medicine*, *74*(4), 377-386. https://doi.org/10.1097/PSY.0b013e3182545d47

## Resurse Online - Design-uri de Cercetare Cvasi-Experimentale

1. Time-Series Study - SAGE Research Methods
2. The Value of Interrupted Time-Series Experiments for Community Intervention Research
3. Design and Analysis of Time Series Experiments
4. DESIGN AND ANALYSIS OF RESEARCH USING TIME SERIES
5. Equivalent_time-samples_design - Equivalent time-samples design Definition and Distinctive Features This is a quasi-experimental research design(see
6. Fostering Critical-Thinking Abilities and Their Effectiveness -- An RN-BSN Program in Taiwan
7. Equivalent_materials_pretest-posttest_design - Equivalent materials pretest posttest design Definition and Distinctive Features It is a
8. https://academicjournals.org/journal/ERR/article-full-text-pdf/01AA00749743.pdf
9. https://www.jstor.org/stable/40317023?seq=1#metadata_info_tab_contents
10. 8.2 Non-Equivalent Groups Designs – Research Methods in Psychology
11. Nonequivalent Groups Design | Research Methods Knowledge Base
12. Nonequivalent Group Design - Reichardt - 2005 - Major Reference Works
13. Counterbalanced Measures Design - Counterbalancing Test Groups
14. Counterbalancing - SAGE Research Methods
15. Counterbalancing in Research
16. Counterbalancing | Topics | Psychology | tutor2u
17. Separate Pre-Post Samples design
18. http://statmath.wu.ac.at/~salzberger/quantmeth/pdfdownloaddirectory/Topic_8_Trochim_chapter10_quasiexperimentaldesign.pdf
19. https://academic.oup.com/cs/article-abstract/29/2/87/394649?redirectedFrom=PDF

20. A separate-sample pretest-post-test design to evaluate a... : Academic Medicine
21. https://socialinquiry.files.wordpress.com/2011/08/lect-7_experiments_on-line.ppt
22. Other Quasi Designs | Research Methods Knowledge Base
23. Cross-sectional time-series experiments: Some suggested statistical analyses.
24. Understanding and using time series analyses in addiction research
25. Multiple Time Series Design
26. What is a multiple time series design?
27. Research Designs
28. Pre-experimental Designs for Research
29. The Concept of Exchangeability in Designing Telecourse Evaluations
30. + Survey and Qualitative Research
31. Regression-Discontinuity Analysis | Research Methods Knowledge Base
32. A Practical Guide to Regression Discontinuity
33. Regression Discontinuity Analysis
34. Data Analysis Notes
35. Modern Regression Discontinuity Analysis

# Design-uri de Cercetare Non-Experimentale

## Studiu Transversal (*Cross-Sectional*)

### Definiţie şi Caracteristici

Studiile transversale sunt preocupate de obţinerea unui eşantion reprezentativ prin analiza unei secţiuni transversale dintr-o populaţie. Toate măsurătorile, în cadrul acestui design, sunt obţinute într-un singur punct temporal, în ciuda faptului că recrutarea participanţilor poate fi de durată (e.g., pe parcursul a doi ani). Totodată, design-ul transversal este adecvat pentru estimarea prevalenţei unui comportament sau a unei tulburări dintr-o populaţie. Deseori, aceste tipuri de studii sunt bazate pe date obţinute din sondaje sau chestionare (Sedgwick, 2014a).

### Avantaje şi Dezavantaje

Ca avantaje, studiile transversale sunt, în general, rapide şi uşor, respectiv ieftin de implementat. Nu există riscul de pierdere a respondenţilor la *follow-up*, deoarece se măsoară o singură dată. Totodată, pot fi înregistraţi mai mulţi factori de risc şi pot fi evaluate mai multe consecinţe. Astfel, studiile transversale pot informa ipoteze ale viitoarelor cercetări mai complexe (e.g., studii pe cohorte) (Sedgwick, 2014a).

Ca dezavantaje, studiile transversale pot fi expuse riscului erorii de participare (engl., "*participation*/ *non-response bias*", momentul în care participanţii care nu răspund sunt diferiţi de participanţii care răspund). În acest sens, prevalenţele riscă să fie subapreciate sau supraapreciate de către cercetători (Prince, 2011). Nu în ultimul rând, cea mai importantă limitare a studiilor transversale este aceea că pot fi inferate doar asocieri, nu şi cauzalităţi (Sedgwick, 2014a).

## Studiu de Tip *Case-Control*

### Definiţie şi Caracteristici

Studiile de tip *case-control* sunt un tip de studii observaţionale. Participanţii sunt recrutaţi în baza statutului unei variabile de interes (e.g., afecţiune, tulburare, boală, etc.). Totodată, unii participanţi au parte de un rezultat în privinţa variabilei de interes (i.e., participanţi din grupul *cases*), alţii nu au parte de un astfel de rezultat (i.e., participanţii din grupul *control*). Condiţiile participanţilor *cases* ar trebui să fie stipulate cât mai specific şi detaliat posibil. Pentru a se asigura faptul că grupul *cases* şi grupul *control* prezintă, majoritar, caracteristici similare, deseori cercetătorii utilizează tehnici de *matching* (Setia, 2016).

### Avantaje şi Dezavantaje

Ca avantaje, studiile de tip *case-control* sunt relativ rapide şi necostisitoare, utile pentru studierea unor consecinţe rare şi ale unor consecinţe cu perioadă lungă de latenţă. Totodată, acestea sunt utile şi în studierea expunerilor multiple la aceeaşi consecinţă, respectiv în studierea asocierii dintre factorii de risc şi consecinţe, în cazul unei epidemii (Setia, 2016).

Ca dezavantaje, studiile de tip *case-control* nu sunt, neapărat, utile în studierea expunerilor rare (i.e., mai util ar putea fi un studiu pe cohorte), sau în studierea consecinţelor multiple (i.e., este un design limitat la o consecinţă specifică). În cadrul acestor studii, nu pot fi estimate incidenţe şi prevalenţe. Câteodată, temporalitatea expunerii şi consecinţa nu sunt stabilite clar într-un astfel de studiu. Nu în ultimul rând, studiile de tip *case-control* pot fi eronate (e.g., eroarea de selecţie şi eroarea de amintire, "*recall bias*") (Setia, 2016).

# Studiu de Tip Înainte-După (*Before-and-After*)

## Definiţie şi Caracteristici

Denumit şi "*pre-post study*", acest tip de studiu măsoară consecinţele într-un grup de participanţi înainte şi după introducerea unui tratament sau a unei intervenţii. Astfel, identificarea schimbărilor se face cu ajutorul unei consecinţe specifice. Acest tip de studiu ajută la explorarea eficienţei tratamentului sau a intervenţiei (Public Health England, 2020). Totodată, studiile de tip înainte-după pot fi "necontrolate" sau "controlate", în funcţie de natura design-ului de cercetare utilizat (Sedgwick, 2014b).

## Avantaje şi Dezavantaje

Ca avantaje, studiile de tip înainte-după ne pot spune ceva despre efectul tratamentului sau al intervenţiei. Totodată, acestea nu necesită selectare sau alocare aleatorie, nu sunt costisitoare, din punctul de vedere al resurselor şi pot fi derulate în timp ce tratamentul sau intervenţia sunt aplicate (Public Health England, 2020).

Ca dezavantaje, studiile de tip înainte-după nu pot spune cu certitudine dacă tratamentul sau intervenţia este cauza reală a unui efect observat. Adiţional, în cazul în care nu există diferenţe înainte şi după administrarea tratamentului sau a intervenţiei, rezultatele devin dificil de interpretat (Public Health England, 2020).

# Studiu Retrospectiv de Tip Istoric cu Grup de Control (*Historical Controls*)

## Definiţie şi Caracteristici

Studiile de tip *historical controls* constau în compararea unui nou tratament sau a unei noi intervenţii, pe un anumit set de subiecţi, cu un tratament vechi sau cu o intervenţie veche, pe un set anterior de subiecţi. Acest tip de design este utilizat în cazuri de limitări etice, în privinţa grupului de control (Baker & Lindeman, 2001).

## Avantaje şi Dezavantaje

Dintre avantaje, amintim costurile reduse ale studiilor de tip *historical controls*, alături de posibilitatea de alocare a multiple resurse grupului experimental (Baker & Lindeman, 2001).

Ca dezavantaje, pot fi amintite inabilitatea de selectare sau alocare aleatorie, inabilitatea de *blinding*, metodologia diferită de diagnosticare (i.e., cunoscută şi sub denumirea de "eroare de

replicare"), efectele specifice tendinţelor temporale (Pandis & Cobourne, 2013), respectiv eroarea de selecţie (Baker & Linderman, 2001).

# Sondaje/ Chestionare

## Definiţie şi Caracteristici

Aceste tipuri de design-uri reprezintă metode de colectare a datelor fie în sens descriptiv, fie în sens predictiv (Goddard III & Villanova, 2005). O caracteristică importantă este aceea că sunt restrânse la un eşantion reprezentativ pentru grupul de care este interesat un cercetător, din motive de practicalitate şi de cost-eficacitate. Sondajele şi chestionarele pot fi transversale, longitudinale (e.g., pe cohorte sau pe tendinţe) sau corelaţionale. Adiţional, acestea dispun de trei metode de colectare a datelor: (1) interviuri faţă-în-faţă; (2) interviuri telefonice; şi (3) chestionare (Mathers, Fox, & Hunn, 2007).

## Avantaje şi Dezavantaje

Printre avantajele sondajelor şi chestionarelor, putem include validitatea internă şi externă, eficacitatea, faptul că pot acoperi eşantioane distribuite geografic, pot avea avantaje etice (e.g., neinvazive) şi sunt flexibile (Mathers et al., 2007).

Dezavantajele sondajelor şi chestionarelor constau în faptul că acestea sunt dependente de cadrul eşantionării, neexplicând de ce oamenii gândesc sau fac ceea ce se poate observa. Adiţional, sondajele de tip interviu au eficacitatea direct proporţională cu abilitatea intervievatorului (Mathers et al., 2007).

# Studiu de Caz de Tip *Case Series*

## Definiţie şi Caracteristici

Cunoscute şi drept studii "*clinical series*", acestea descriu experienţele unui grup restrâns de subiecţi. Se poate afirma, în acest sens, că este o variaţie a unui studiu de caz de tip *single case report* (discutat în subcapitolul următor). Studiile de caz de tip *case series* reprezintă design-uri de cercetare observaţionale, descriptive, ce pot fi derulate atât retrospectiv, cât şi prospectiv. În acest sens, cercetătorul nu manipulează intervenţia sau tratamentul, ci doar descrie ce se întâmplă în cursul normal al acesteia sau al acestuia, prezentând subiecţii într-un mediu clinic natural (pentru mai multe detalii, vezi El-Gilany, 2018a).

## Avantaje şi Dezavantaje

Ca avantaje, studiile de tip *case series* aduc valoare în cercetarea epidemiologică. În acest sens, sunt utile pentru generarea de ipoteze (i.e., nu pot fi trase concluzii cu privire la etiologie), utilizează doar informaţii despre cazuri (i.e., nu este nevoie de un grup de control), derularea studiului este de scurtă durată şi acesta nu interferează cu procesul de decizie privind intervenţia sau tratamentul (pentru mai multe detalii, vezi El-Gilany, 2018a).

Dintre dezavantaje, amintim faptul că studiile de tip *case series* nu pot investiga relaţii de tip cauză-efect (i.e., nu poate fi inferată cauzalitatea), nu pot evalua frecvenţa bolii, a tulburării sau a

afecţiunii, sunt pasibile de eroare de selecţie (engl., "*selection bias*"). Adiţional, o anumită consecinţă poate fi cauzată de şansă, nu neapărat de o boală, tulburare sau afecţiune. Nu în ultimul rând, studiile de tip *case series* au o validitate internă scăzută, din cauza lipsei grupului de control (pentru mai multe detalii, vezi El-Gilany, 2018a).

# Studiu de Caz de Tip *Case Report*

## Definiţie şi Caracteristici

Studiul de tip *case report* este un tip specific de design de cercetare ce raportează un aspect pe unul sau maxim doi subiecţi. În acest sens, este utilizat pentru a descrie situaţii excepţionale ce reprezintă o provocare pentru teoreticieni şi practicieni deopotrivă. Aşadar, se numără printre cele mai puţin publicate forme de cercetare ştiinţifică. Cu toate acestea, poate aduce noi cunoştinţe şi valori ce ţin de informarea practicienilor. Un aspect important de reţinut este acela că trebuie să fie bine făcută distincţia dintre studiile de tip *case series*, cele de tip *case studies*, studiile clinice şi cele de tip *case report form* (pentru mai multe detalii, vezi El-Gilany, 2018b).

## Avantaje şi Dezavantaje

Printre avantajele studiilor de tip *case reports* se numără descrierea observaţiilor ştiinţifice şi clinice ce ar putea fi omise de o "treaptă mai înaltă" a ierarhiei evidenţelor (e.g., meta-analize, studii clinice controlate, etc.; acest aspect va fi explorat în cadrul subcapitolului materialului actual dedicat ierarhiei evidenţelor). Totodată, acest tip de studii reprezintă propuneri valide pentru studii cu design-uri mai puternice, fiind capabil de generarea de ipoteze. Deseori, sunt utilizate atât în demersuri de învăţare, cât şi în demersuri de cercetare (pentru mai multe detalii, vezi El-Gilany, 2018b).

Printre dezavantajele studiilor de tip *case reports* se numără nivelul scăzut de evidenţă, potenţial cauzat de erori de observare, faptul că rezultatele nu pot fi generalizate, faptul că acestea nu prezintă date cu ajutorul cărora să se poată calcula rate de prevalenţă ale anumitor boli, tulburări sau afecţiuni, nu prezintă o bună validitate ştiinţifică (i.e., prin lipsa grupului de control) şi sunt articole infrecvent citate, având potenţialul de a scădea factorul de impact al unui jurnal (pentru mai multe detalii, vezi El-Gilany, 2018b).

## Sumar - Design-uri de Cercetare Non-Experimentale

- Studiile transversale reprezintă observaţii ale unui grup într-un singur moment de timp. Acestea sunt utile pentru calcularea ratelor de prevalenţă şi pentru colectarea rapidă a datelor preliminare (Thompson & Panacek, 2007).
- Studiile de tip *case-control* sunt utile pentru testarea unei potenţiale cauze a unei afecţiuni. Aceste studii sunt structurate de la efect înspre cauză şi conţin grup de control (Thompson & Panacek, 2007).
- Studiile de tip înainte-după cuantifică schimbarea într-o formă de terapie sau de mediu, între două măsurări (Thompson & Panacek, 2007).
- Studiile de tip *historical controls* sunt similare cu cele înainte-după. Acestea nu necesită o perioadă absolută axată pe "schimbare". Studiile de tip *historical controls* sunt utilizate atunci când un grup de control nu poate fi inclus (i.e., din considerente etice sau de altă natură). În acest sens, foşti respondenţi (i.e., din date istorice) formează grupul de control (Thompson & Panacek, 2007).
- Sondajele şi chestionarele adresează în mod direct întrebările cercetării şi "adună" răspunsurile. Tocmai din acest punct de vedere, există un potenţial ridicat de erori, în cazul în care nu sunt corect întocmite (Thompson & Panacek, 2007).
- Studiile de tip *case series* cuprind descrierea unei serii de respondenţi cu aceeaşi caracteristică definită. Un astfel de tip de studiu poate fi derulat fie cu ajutorul unui grup consecutiv, fie selectiv, neprezentând grup de control (Thompson & Panacek, 2007).
- Studiile de tip *case report* sunt preocupate de descrierea unui singur caz, cu scopul de a raporta noi descoperiri. Astfel, acestea constituie un demers unic de educare şi informare (Thompson & Panacek, 2007).

## Resurse Online - Design-uri de Cercetare Non-Experimentale

1. Methodology Series Module 3: Cross-sectional Studies
2. Cross-Sectional Study | Definitions, Uses & Examples
3. Cross Sectional Study - an overview
4. Epidemiology in Practice: Case-Control Studies
5. Case Control - Study Design 101
6. Definition of case-control study - NCI Dictionary of Cancer Terms
7. Case-Control Studies - an overview
8. Statistical inference in matched case–control studies of recurrent events
9. Quality improvement studies – pitfalls of the before and after study design
10. Before-and-after studies • LITFL • CCC Research
11. Before-and-after design: A simple evaluation design
12. Bias in Before–After Studies: Narrative Overview for Anesthesiologists
13. The Limitations of Before and After Designs
14. Historical control | definition of historical control
15. Use of historical control data for assessing treatment effects in clinical trials
16. Historical Controls?
17. Historical Controls for Clinical Trials
18. A roadmap to using historical controls in clinical trials – by Drug Information Association Adaptive Design Scientific Working Group (DIA-ADSWG)
19. Survey vs Questionnaire: What's the difference?

20. 24 Questionnaire Examples, Questions, & Templates to Survey Your Clients
21. Psychological Assessment Scales And Measures
22. Questionnaires for Researchers
23. Researchcentral
24. Clarifying the distinction between case series and cohort studies in systematic reviews of comparative studies: potential impact on body of evidence and workload
25. Clinical “case series”: a concept analysis
26. How to Design a Good Case Series
27. © Joanna Briggs Institute 2017 Critical Appraisal Checklist for Case
28. Retrospective Study: Case-Control and Case-Series
29. How to write a patient case report
30. Journal of Medical Case Reports | Home page
31. Guidelines To Writing A Clinical Case Report
32. Case report
33. Guidelines for Submitting Case Reports
34. Case Studies
35. Modele de cercetare non-experimentale 1 Introducere
36. Overview of Nonexperimental Research – Research Methods in Psychology
37. http://psychology.iresearchnet.com/social-psychology/social-psychology-research-methods/nonexperimental-designs/
38. Chapter 7: Nonexperimental Research

# Alte Design-uri de Cercetare

## Studii Corelaţionale

### Definiţie şi Caracteristici

Studiile corelaţionale sunt studii descriptive, similare cu studiile transversale, despre care am discutat anterior. Scopul principal al studiilor corelaţionale este acela de a explora asocieri între variabile. În plus, acestea pot fi implementate cu uşurinţă şi pot acumula cantităţi relativ mari de date, relativ rapid (Thompson & Panacek, 2007).

### Limitări

Având în vedere că studiile corelaţionale urmăresc asocierile dintre variabile, este important de menţionat că o asociere nu marchează, neapărat, faptul că variabilele respective sunt cu adevărat asociate (i.e., asocierea poate fi datorată şansei). Ba chiar nici evidenţele puternice ale asocierilor dintre variabile nu reprezintă, neapărat, cauzalitate. Totodată, relaţiile temporale dintre variabile sunt, deseori, incerte (i.e., dileme de tip "Cine a fost primul? Oul sau găina?") (Thompson & Panacek, 2007).

## Meta-Analize

### Definiţie şi Caracteristici

Meta-analizele pot fi definite ca modele de cercetare ce nu implică o nouă colectare de date, ci mai degrabă o reanalizare a datelor anterior colectate şi raportate de către alţi cercetători. Aceste modele de cercetare presupun ca analizele statistice ale rezultatelor din două sau mai multe studii independente să fie analizate cu scopul de integrare a rezultatelor şi de dezvoltare a unor concluzii generale. În particular, meta-analizele sunt utile atunci când numărul de studii derulate corespunzător este scăzut, în literatura de specialitate. Astfel, o meta-analiză poate sistematiza rezultatele unor studii mici, oferind, astfel, concluzii generale optimizate (Thompson & Panacek, 2007).

### Limitări

Firește, chiar şi meta-analizele pot veni cu limitări, în funcţie de criteriile robuste de incluziune, în cazul în care acestea nu sunt suficient de robuste. Totodată, calitatea globală a meta-analizelor este pe atât de robustă pe cât este şi calitatea studiilor incluse în acestea (Thompson & Panacek, 2007).

# Revizuiri Sistematice (*Systematic Reviews*)

## Definiţie şi Caracteristici

Revizuirile sistematice sunt studii ce implică o recenzie (denumită şi "revizuire") cât mai obiectivă cu putinţă a corpului de literatură pe un anumit subiect, cu scopul de identificare a concluziilor şi recomandărilor. În cadrul revizuirilor sistematice, nu sunt utilizate proceduri statistice meta-analitice pentru sumarizarea evidenţelor. Aceste studii sunt utilizate cu scopul de popularizare a curentului *evidence-based* (Thompson & Panacek, 2007).

## Limitări

Ca limitări ale revizuirilor sistematice, amintim gradul de validitate ştiinţifică (i.e., acesta poate varia, în funcţie de procesul de selecţie al articolelor utilizat) şi subiectivitatea (i.e., fapt dependent de opiniilor autorului sau ale autorilor) (Thompson & Panacek, 2007).

# Studii Cost-Beneficiu (*Cost-Benefit*)

## Definiţie şi Caracteristici

Studiile cost-beneficiu reprezintă forme de evaluare economică în care costurile unei anumite intervenţii sunt comparate cu beneficiile intervenţiei propriu-zise. În acest sens, un număr de asumpţii este deseori inclus în analiză (e.g., "Care este valoarea unui an suplimentar, în privinţa acestei condiţii?") (Thompson & Panacek, 2007).

## Limitări

Limitările studiilor cost-beneficiu constau în identificarea naturii, a intervalelor şi a motivelor din spatele asumpţiilor anterior menţionate (i.e., acest fapt poate reprezenta o sursă de eroare, pentru modelele de analiză propuse) (Thompson & Panacek, 2007).

# Studii Cost-Eficacitate (*Cost-Effectiveness*)

## Definiţie şi Caracteristici

Similare cu studiile de tip cost-beneficiu, studiile cost-eficacitate compară programe, terapii sau intervenţii alternative, în termeni de costuri generale per grad de efect clinic (e.g., terapia X poate creşte cu 1% bunăstarea autoraportată a subiecţilor, într-un caz specific). Acest tip de studii necesită efort de grup (e.g., grupuri de clinicieni, biostatisticieni şi, câteodată, economişti) (Thompson & Panacek, 2007).

## Limitări

Limitările acestui tip de studii constau chiar în faptul că metodologia specifică studiilor cost-eficienţă (i.e., şi chiar şi cost-beneficiu) nu este una standardizată. În acest sens, acest tip de

studii prezintă variabilitate mare, fiind un model de cercetare avansat, ce necesită un anumit nivel de expertiză (Thompson & Panacek, 2007).

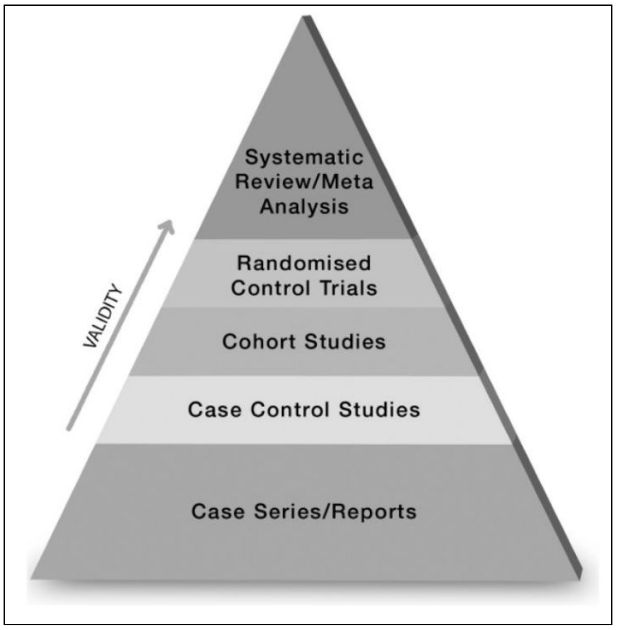
## Resurse Online - Alte Design-uri de Cercetare

1. How Correlational Studies Are Used in Psychology
2. An Example Of A Correlational Study
3. Meta-Analysis - an overview
4. https://www.meta-analysis.com/pages/why_do.php?cart=BRED4802813
5. Systematic Reviews | Home page
6. What is a systematic review? (pdf)
7. Cost Benefit Analysis: An Expert Guide
8. Cost-Benefit Analysis (CBA)
9. Cost-Effectiveness Analysis
10. Cost-effectiveness analysis for health interventions

# Ierarhia Evidenţelor

## Accepţia Clasică

Accepţia clasică a ierarhiei evidenţelor poate fi caracterizată ca un model structurat ierarhic, rigid şi suprasimplificat. Schema generală a ierarhiei evidenţelor este una destul de cunoscută. În acest sens, conform articolului lui Murad, Asi, Alsawas şi Alahdab (2016, p. 126), schema va fi ilustrată, în continuare.




## Scurte Definiţii

Pentru o scurtă recapitulare a modelelor de cercetare şi a tipurilor de studii ierarhizate în cadrul "piramidei" evidenţelor, se va face o scurtă trecere în revistă a definiţiilor acestora, după cum urmează.

### *Case Series* şi *Case Reports*

Un studiu de tip *case report* reprezintă un raport detaliat al diagnosticului, tratamentului sau intervenţiei, răspunsului la tratament sau intervenţie şi *follow-up* după tratament sau intervenţie, pentru un singur subiect. Astfel, un studiu de tip *case series* reprezintă un grup de studii de tip *case report* care implică tratament sau intervenţie similar(ă) pentru subiecţii analizaţi (Boston University Medical Campus, n.d.).

### *Case Control Studies*

Studiile de tip *case control* reprezintă studii observaţionale cu grup de control, în care participanţii sunt selectaţi în baza statutului variabilei (i.e., consecinţei) de interes pentru cercetător (Setia, 2016).

### *Cohort Studies*

Studiile pe cohorte sunt studii în care unul sau mai multe eşantioane (denumite "cohorte") sunt urmărite prospectiv, cu scopul de a determina ce caracteristici iniţiale ale participanţilor sunt asociate cu o anumită consecinţă (e.g., boli, tulburări, afecţiuni, etc.) (Himmelfarb Health Sciences Library, n.d.).

### *Randomized Controlled Trials* (*RCTs*)

Studiile clinice controlate sunt studii prospective ce măsoară eficacitatea unei noi intervenţii sau a unui nou tratament (Hariton & Locascio, 2018).

### *Systematic Review* şi *Meta-Analysis*

Revizuirile sistematice sunt cercetări ce implică un plan detaliat şi comprehensiv şi o strategie de căutare propuse *a priori* (i.e., înainte de a începe cercetarea propriu-zisă), în sensul reducerii cantităţii de erori prin identificarea, evaluarea şi sintetizarea tuturor studiilor relevante, pe o anumită temă (Uman, 2011). Meta-analizele pot fi privite drept componente ale revizuirilor sistematice, ce presupun utilizarea de proceduri statistice pentru sintetizarea datelor din multiple studii într-o singură estimare cantitativă sau mărime a efectului (Pattincrew & Roberts, 2006, în Uman, 2011).

## Ediţia Revizuită

Astfel, ediţia revizuită a ierarhiei evidenţelor propune un model dinamic, în care treptele superioare de evidenţă reprezintă o lentilă prin care pot fi privite celelalte forme de evidenţă. Pentru integrarea acestei informaţii, schema generală a modelului, conform articolului lui Murad şi al colaboratorilor săi (2016, p. 126), este propusă, în continuare.

## Resurse Online - Ierarhia Evidențelor

1. Hierarchy of Evidence and Study Design - OHSU Evidence-Based Practice Course for Interprofessional Clinical Teams
2. Pyramids are guides not rules: the evolution of the evidence pyramid
3. Academic Guides: Evidence-Based Practice Research: Levels of Evidence Pyramid
4. UC Library Guides: Evidence-Based Practice in Health: Hierarchy of Evidence
5. Hierarchy of evidence
6. OCEBM Levels of Evidence
7. https://www.youtube.com/watch?v=-omKRqfQaY4&t=7s

# Statistică

# Fidelitatea

## Definiţie şi Tipuri de Fidelitate

Fidelitatea reprezintă abilitatea de replicare a rezultatelor într-o modalitate consistentă în timp şi în spaţiu (Souza, Alexandre, & Guirardello, 2017).

Există 4 tipuri de fidelitate, după cum urmează:

- Fidelitatea inter-evaluator.
- Fidelitatea test-retest (intra-evaluator).
- Fidelitatea de tip forme paralele (forme echivalente).
- Fidelitatea de tip consistenţă internă.

### Fidelitatea Inter-Evaluator

#### Definiţie şi Caracteristici

Fidelitatea inter-evaluator poate fi înţeleasă drept măsura în care doi sau mai mulţi evaluatori sunt de acord în privinţa măsurătorilor efectuate. Fidelitatea inter-evaluator adresează subiectul consistenţei implementării unui sistem de scorare. În acest sens, o fidelitate inter-evaluator mare indică un grad ridicat de acord între cei doi evaluatori, în timp ce o fidelitate inter-evaluator mică indică un grad scăzut de acord între cei doi evaluatori (Lange, 2011).

Pentru o înţelegere aplicată a conceptului de fidelitate inter-evaluator, se va face apel la un articol (Lobbestael, Leurgans, & Arntz, 2011) din literatura de specialitate.

#### Statistici Uzuale

Pentru cuantificarea fidelităţii inter-evaluator, există o serie de statistici uzuale utilizate, listate şi explicate după cum urmează:

- Procentajul de acord.
- *Kappa*-ul lui Cohen (*Cohen's Kappa*).
- Corelaţia produs-moment (*product-moment correlation*, coeficientul $\rho$ de corelaţie al lui Spearman).
- Coeficientul de corelaţie intra-clasă (*intraclass correlation coefficient*).

#### Procentajul de Acord

Această metrică poate fi înţeleasă prin intermediul formulei: (numărul de acorduri privind observaţiile/ numărul total de observaţii)*100 (Lange, 2011).

#### *Kappa*-ul lui Cohen

Indicele *kappa* al lui Cohen este un coeficient standardizat ce ia valori între -1 şi +1. În acest sens, McHugh (2012, p. 279) propune un ghid de interpretare a valorilor coeficientului, ce va fi ilustrat în continuare.

| Value of Kappa | Level of Agreement | % of Data that are Reliable |
|---|---|---|
| 0–.20 | None | 0–4% |
| .21–.39 | Minimal | 4–15% |
| .40–.59 | Weak | 15–35% |
| .60–.79 | Moderate | 35–63% |
| .80–.90 | Strong | 64–81% |
| Above.90 | Almost Perfect | 82–100% |

### Corelaţia Produs-Moment

Cunoscut şi drept coeficientul $\rho$ de corelaţie al lui Spearman, această metrică indică măsura gradului de dependenţă dintre două variabile selectate aleatoriu (Nishisato & Torii, 1971).

### Coeficientul de Corelaţie Intra-Clasă

Această metrică implică o procedură statistică bazată pe analiza varianţei şi pe estimarea componentei acesteia. În sens univariat, rata de varianţă poate fi interpretată drept coeficient de corelaţie. În sens bivariat (i.e., varianţa inter-evaluator este eliminată), se recomandă precauţie în interpretare (Bartko, 1966).

## Fidelitatea Test-Retest/ Intra-Evaluator

Acest tip de fidelitate presupune măsurarea subiecţilor în două ocazii diferite, cu acelşai instrument, fapt urmat de indicarea corelaţiei dintre măsurători (Collins, 2007). Totodată, fidelitatea test-retest poate fi înţeleasă şi drept gradul în care scorurile obţinute de la aceiaşi participanţi rămân consistente pe o perioadă redusă de timp, în care competenţele şi/ sau problemele respondenţilor nu se modifică semnificativ (McConaughy, 2001).

Astfel, fidelitatea intra-evaluator necesită două administrări ale aceluiaşi instrument, într-o perioadă de timp în care nu au survenit schimbări în conceptul-ţintă măsurat (Beaton, Boers, & Tugwell, 2013).

Pentru o înţelegere aplicată a conceptului de fidelitate test-retest, se va face apel la un articol (Backhaus, Junghanns, Broocks, Riemann, & Hohagen, 2002) din literatura de specialitate.

## Fidelitatea de Tip Forme Paralele/ Forme Echivalente

Acest tip de fidelitate presupune subseturi diferite ale aceluiaşi set de itemi (i.e., din cadrul aceleiaşi scale) ce surprind atribute similare, cu precizie similară. Comparativ cu fidelitatea test-retest, fidelitatea de tip forme echivalente previne eforturile sistematice de corespondenţă între prima măsurare a unui atribut şi cea de-a doua. Astfel, nu există riscul de a surveni erori de memorare a testelor sau de obişnuinţă a respondenţilor cu testele. Totodată, fidelitatea de tip forme paralele conferă posibilitatea măsurătorilor repetate ale aceluiaşi atribut, independent de prima măsurare (Hilger & Beauducel, 2017).

Pentru o înţelegere aplicată a conceptului de fidelitate de tip forme paralele, se va face apel la un articol (Brandt, 1991) din literatura de specialitate.

## Fidelitatea de Tip Consistenţă Internă

### Definiţie şi Caracteristici

Fidelitatea de tip consistenţă internă reflectă măsura în care itemii unui instrument măsoară anumite aspecte ale aceleiaşi caracteristici sau ale aceluiaşi construct (Revicki, 2014). Acest tip de fidelitate estimează cât de mult ar varia scorurile totale ale unui test, în momentul în care ar fi utilizaţi itemi relativ diferiţi (Salkind, 2010c).

### Statistici Uzuale

Pentru cuantificarea fidelităţii de tip consistenţă internă, există o serie de statistici uzuale utilizate, după cum urmează:

- $\alpha$-ul lui Cronbach (*Cronbach's Alpha*).
- Corelaţia medie inter-itemi.
- Fidelitatea *Split Half*.
- *Kuder-Richardson 20* (*KR-20*).

Pentru o înţelegere aplicată a conceptului de fidelitate de tip forme paralele, se va face apel la un articol (McCrae, Kurtz, Yamagata, & Terracciano, 2011) din literatura de specialitate.

#### $\alpha$-ul lui Cronbach

Această metrică este una dintre cele mai des întâlnite şi utilizate metrici din cercetarea aplicată în domeniul psihologiei. Poate fi înţeles drept o măsură a consistenţei interne a unui test, ce ia valori între 0 şi 1. Este important de menţionat că un coeficient $\alpha$ nu arată, întotdeauna, o consistenţă internă ridicată. Adiţional, este cunoscut că indicele $\alpha$ este afectat de lungimea unui test (Tavakol & Dennick, 2011).

#### Corelaţia Medie Inter-Itemi

Această metrică reprezintă, pur şi simplu, media corelaţiilor dintre toate perechile de itemi (Glen, 2018).

#### Fidelitatea *Split Half*

Acest tip de fidelitate implică înjumătăţirea unui test şi corelarea scorurilor respondenţilor pe cele două jumătăţi de test. Astfel, necesită o singură administrare a unui test, reprezentând o alternativă convenientă a metodei test-retest şi a fidelităţii de tip forme echivalente. Fidelitatea *Split Half* a fost utilizată, majoritar, înaintea descoperirii indicelui $\alpha$ (Zhu & Lowe, 2018).

#### *Kuder-Richardson 20* (*KR-20*)

Această formulă de fidelitate este utilizată pentru măsurătorile cu variante dihotomice de răspuns (e.g., cu 0 şi 1, cu da şi nu, etc.) (Nugroho, Warnars, Heriyadi, & Tanutama, 2019).

## Sumar - Fidelitatea

- Fidelitatea presupune repetarea consistentă a rezultatelor.
- Există 4 forme de fidelitate.
- Fidelitatea inter-evaluator reprezintă măsura acordului dintre doi sau mai mulţi utilizatori. Cele mai des întâlnite metrici ale fidelităţii inter-evaluator sunt: procentajul de acord, *kappa*-ul lui Cohen, corelaţia produs-moment şi coeficientul de corelaţie intra-clasă.
- Fidelitatea test-retest presupune măsurarea aceloraşi participanţi, în două momente diferite de timp, cu aceleaşi instrumente, de către aceiaşi evaluatori.
- Fidelitatea de tip forme paralele presupune subseturi paralele/ echivalente ale aceluiaşi test, în cadrul unor măsurători repetate independente.
- Fidelitatea de tip consistenţă internă reprezintă corelaţia testului cu el însuşi. Cele mai des întâlnite metrici ale fidelităţii de tip consistenţă internă sunt: *alfa*-ul lui Cronbach, corelaţia medie inter-itemi, fidelitatea *Split Half* şi indicele *KR-20*.

## Bibliografie Selectivă - Fidelitatea


Backhaus, J., Junghanns, K., Broocks, A., Riemann, D., & Hohagen, F. (2002). Test–retest reliability and validity of the Pittsburgh Sleep Quality Index in primary insomnia. *Journal of Psychosomatic Research*, *53*(3), 737-740. https://doi.org/10.1016/S0022-3999(02)00330-6

Brandt, J. (1991). The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. *The Clinical Neuropsychologist*, *5*(2), 125-142. https://doi.org/10.1080/13854049108403297

Lobbestael, J., Leurgans, M., & Arntz, A. (2011). Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I disorders (SCID I) and Axis II disorders (SCID II). *Clinical Psychology & Psychotherapy*, *18*(1), 75-79. https://doi.org/10.1002/cpp.693

McCrae, R. R., Kurtz, J. E., Yamagata, S., & Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, *15*(1), 28-50. https://doi.org/10.1177%2F1088868310366253


## Resurse Online - Fidelitatea

1. Reliability vs Validity in Research | Differences, Types and Examples
2. The 4 Types of Reliability | Definitions, Examples, Methods
3. What is Reliability?
4. Reliability | Psychology | tutor2u
5. Reliability and Consistency in Psychometrics
6. 4.2 Reliability and Validity of Measurement – Research Methods in Psychology
7. Reliability and Validity
8. Reliability of A Test « A Primer on Psychology

# Analiza de Putere

## Tipuri de Erori Statistice

Tipurile de erori statistice vor fi discutate, în acord cu menţiunile făcute de Kaur şi Stoltzfus (2017) şi de Ottenbacher (1992), listate în tabelul următor.

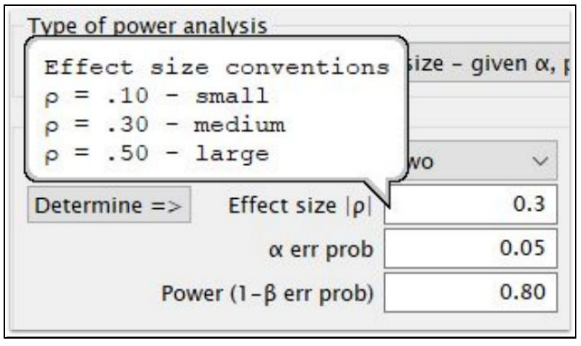
| Eroarea de Tip I | Eroarea de Tip II |
|---|---|
| • Respingerea ipotezei de nul.<br>• Ipoteza de nul este adevărată. | • Acceptarea ipotezei de nul.<br>• Ipoteza de nul este falsă. |
| **Eroarea de Tip III** | **Eroarea de Tip IV** |
| • Respingerea corectă a ipotezei de nul.<br>• Raţionament eronat. | • Respingerea corectă a ipotezei de nul.<br>• Interpretare eronată (vezi fie subcapitolul dedicat validităţii concluziei statistice, din cadrul acestui material, fie "Validitatea Concluziei Statistice", p. 9, https://osf.io/d29pr/). |

## Definiţie şi Caracteristici

Puterea (*1-β*) reprezintă probabilitatea ca un efect semnificativ să poată fi observat, atunci când un efect real chiar există. Totodată, poate fi înţeleasă şi în termeni de probabilitate de evitare a erorii de tip II (Uttley, 2019). Conform lui Uttley (2019), elementele unei analize de putere sunt:

- Setarea pragului *α* (i.e., de obicei, la valoarea .05).
- Setarea puterii vizate în cadrul testului (i.e., de obicei, la valoarea de .80, conform lui Cohen, 1992, în Uttley, 2019).
- Anticiparea mărimii efectului (i.e., prin analiza prospectivă).

Fireşte, analiza de putere este strâns legată de conceptul de mărime a efectului. În acest sens, o captură de ecran a software-ului *G*Power*, versiunea 3.1.9.2 (Faul, Erdfelder, Lang, & Buchner, 2007; Faul, Erdfelder, Buchner, & Lang, 2009) va fi anexată, în continuare, pentru listarea convenţiilor de interpretare a mărimii efectului.



## Factori ce Pot Influenţa Puterea Unui Test

Dintre factorii ce pot influenţa puterea unui test, amintim: pragul de semnificaţie statistică, mărimea efectului relaţiilor de interes, mărimea eşantionului (i.e., cu cât creşte dimensiunea eşantionului, cu atât creşte puterea testului, efectele cu magnitudine mai scăzută devenind mai uşor detectabile), tipul de test utilizat (i.e., testele neparametrice au mai puţină putere decât testele parametrice), respectiv tipul de design utilizat (i.e., design-urile *within-subjects*, cu aceiaşi subiecţi în aceleaşi condiţii, au putere mai mare decât design-urile *between-subjects*, cu subiecţi diferiţi per condiţie) (Uttley, 2019).

## Analiza Prospectivă (*A Priori*)

Analiza de putere *a priori* poate fi înţeleasă ca estimarea dimensiunii eşantionului în baza ratelor maxim tolerabile de erori de tip I şi II şi a mărimii minime a efectului (i.e., semnificativă clinic, practic sau teoretic). Acest tip de analiză de putere implică un calcul al mărimii eşantionului necesar pentru detectarea unui efect cu o anumită magnitudine, la puterea dorită. Cu toate acestea, analiza prospectivă dă dovadă de probabilitate condiţională, având potenţiale interpretări eronate. În acest sens, factorii ce pot influenţa analiza prospectivă sunt erorile de tip I şi II şi mărimea efectului (Kyonka, 2019).

## Analiza Retrospectivă (*Post Hoc*)

Denumită şi "putere observată" (engl., "*observed power*"), analiza *post hoc* presupune utilizarea unui test statistic deja derulat, pentru detectarea unei mărimi a efectului egală cu cea obţinută anterior. Aşadar, o astfel de analiză nu indică, neapărat, puterea reală în detectarea semnificaţiei statistice. În general, analizele retrospective de putere nu asigură rezultate precise (Kyonka, 2019; Zhang et al., 2019).

## Resurse Online - Analiza de Putere

1. https://egap.shinyapps.io/Power_Calculator/
2. Introduction to Power Analysis
3. Power Analysis
4. A Gentle Introduction to Statistical Power and Power Analysis in Python
5. A Researcher's Guide to Power Analysis
6. Estimating a Good Sample Size for Your Study Using Power Analysis
7. Power analysis
8. Shiny - MOTE: An Effect Size Calculator
9. Effect size converter

# Statistici Descriptive Univariate

## Tendinţa Centrală a Distribuţiei

Indicatorii tendinţei centrale a distribuţiei sunt:

- Media (*Mean/ Average*).
- Mediana (*Median*).
- Modul (*Mode*).

### Media (*Mean/ Average*)

Media este cea mai utilizată măsură a tendinţei centrale a unei distribuţii, reprezentând valoarea medie a unui grup de numere (Sykes, Gani, & Vally, 2016). Formula generală a mediei este:

$$(x_1 + x_2 + \ldots + x_n)/ n$$

unde
x1, x2, …, xn - Elemente.
n - Numărul total de elemente.

### Mediana (*Median*)

Mediana poate fi înţeleasă în termeni de scor ce înjumătăţeşte distribuţia. În acest sens, mai este denumită şi "percentila 50". Interpretarea scorurilor, în funcţie de mediană, se face astfel: jumătate din scoruri se află peste mediană, în timp ce cealaltă jumătate din scoruri se află sub mediană (Sykes et al., 2016). Formula generală a medianei este:

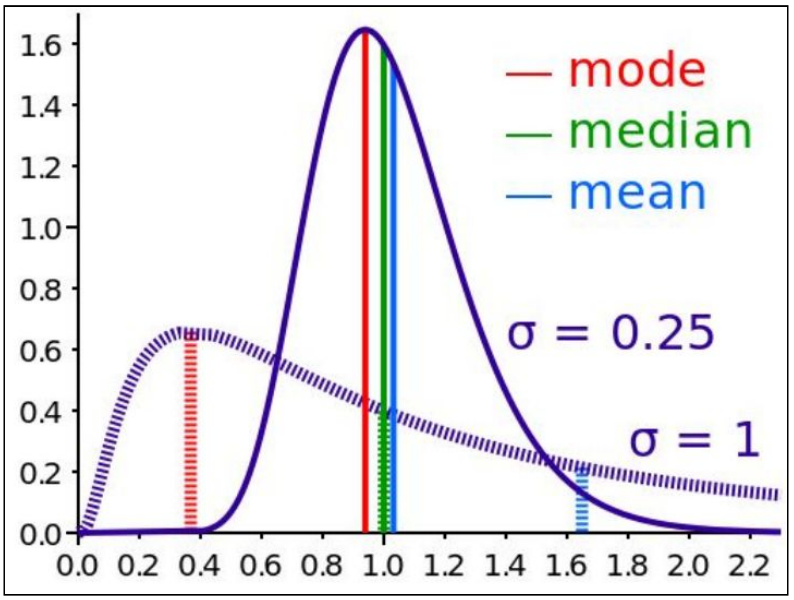
$$(N + 1)/ 2$$

unde
N - Numărul de elemente

### Modul (*Mode*)

Modul reprezintă scorul cu cea mai mare frecvenţă dintr-o distribuţie. Astfel, punctul sau valoarea lui X corespunde celui mai înalt punct din distribuţie. În cazul în care cea mai mare frecvenţă corespunde mai multor puncte (i.e., mai multe valori ale lui X care să corespundă aceluiaşi cel mai înalt punct al distribuţiei), distribuţia se va numi "multimodală" (Sykes et al., 2016).

Pentru înţelegerea aprofundată a conceptelor de medie, mediană şi mod se va utiliza figura propusă de Constantin şi Tresset (n.d., p. 9), după cum urmează.




# Dispersia Distribuţiei

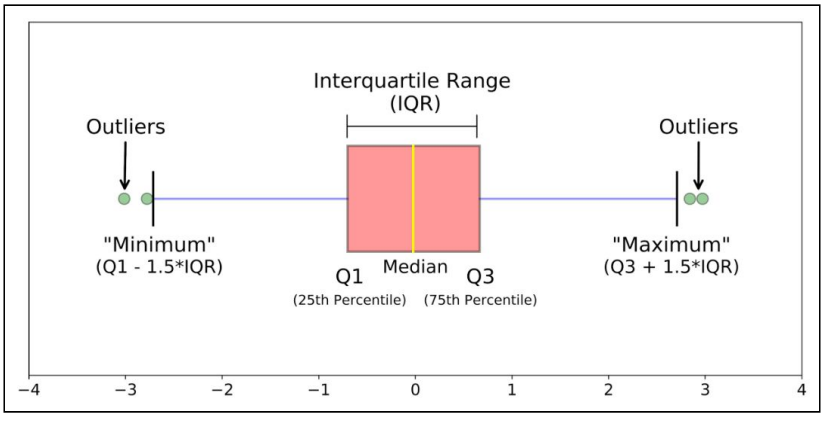
În termeni de dispersie a distribuţiei, indicatorii acesteia sunt:

- Intervalul (*Range*).
- Quartilele (*Quartiles*).
- Abaterea standard (*Standard deviation*).
- Varianţa (*Variance*).

## Intervalul (*Range*)

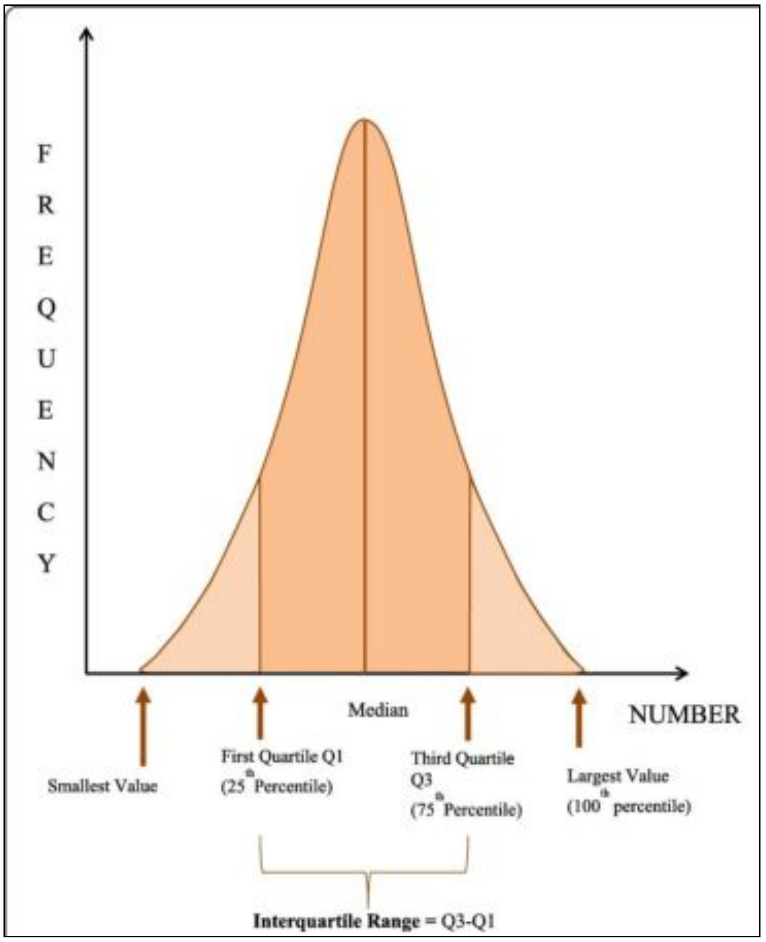
Intervalul reprezintă diferenţa dintre cel mai mare şi cel mai mic scor ale unei distribuţii. Acest indicator nu este des utilizat, deoarece se bazează exclusiv pe valorile extreme ale distribuţiei. În acest sens, nu reflectă un model autentic de variaţie a scorurilor din cadrul unei distribuţii (Sykes et al., 2016).

O aplicaţie utilă a indicatorului discutat este abaterea interquartilă (engl., *Interquartile Range*, *IQR*). Aceasta reprezintă o măsură a dispersiei proporţiei de 50% din scorurile medii ale unei distribuţii. Poate fi definită ca diferenţa dintre percentila 75 şi percentila 25. Metoda sa grafică, ilustrată în continuare (Galarnyk, 2018), este cunoscută sub denumirea de "*boxplot*" (Sykes et al., 2016).



## Quartilele (*Quartiles*)

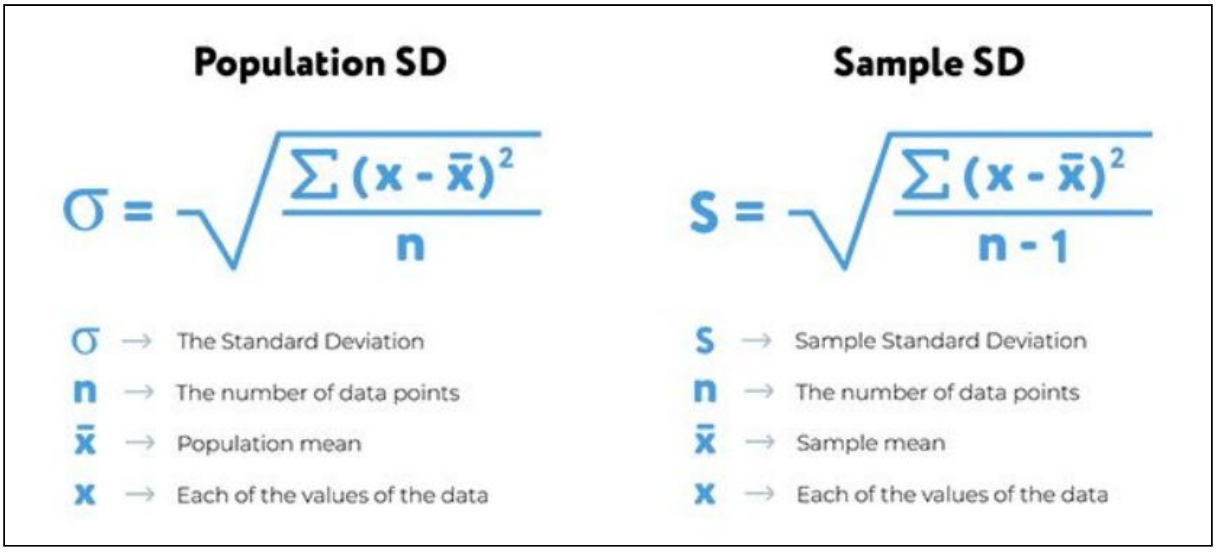
Quartilele reprezintă puncte ce separă setul de date în patru grupuri de dimensiuni egale (Sykes et al., 2016). Pentru vizualizarea conceptului de quartile, se va face apel la figura propusă de Singh şi Oorkavalan (2018, p. 4), după cum urmează în continuare.




## Abaterea Standard (*Standard Deviation*)

În termeni matematici, abaterea standard reprezintă rădăcina pătrată pozitivă a varianţei. Aceasta poate avea două notaţii, în funcţie de întrebuinţare, astfel: (1) $s$, pentru eşantion; şi (2) $\sigma$, pentru populaţie. Simplificat, abaterea standard arată câtă variaţie este într-un anumit grup de valori, măsurând abaterea sau diferenţa de la media grupului de valori. În acest sens, o abatere standard mică poate indica faptul că valorile setului de date se află în apropierea mediei, în timp ce o abatere standard mare poate indica faptul că datele setului sunt distribuite într-un interval mare de valori. În cazul măsurării mai multor constructe, suma tuturor abaterilor standard este egală cu 0 (Sykes et al., 2016).

Formula abaterii standard va fi ilustrată în continuare, cu ajutorul unei figuri (Standard Deviation Calculator, n.d.).


Population SD
$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$
σ → The Standard Deviation
n → The number of data points
$\bar{x}$ → Population mean
x → Each of the values of the data
Sample SD
$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$
S → Sample Standard Deviation
n → The number of data points
$\bar{x}$ → Sample mean
x → Each of the values of the data

### Varianţa (*Variance*)

Nu în ultimul rând, varianţa reprezintă indicatorul bazat pe abaterile standard ale scorurilor individuale de la media distribuţiei. Pentru a evita problematica însumării abaterilor standard, varianţa este bazată pe pătratele abaterilor. Astfel, se păstrează distanţa dintre scorurile distribuţiei, în timp ce scorurile negative sunt eliminate (Sykes et al., 2016). Formula generală a varianţei este:

$$s^2 = \Sigma(\chi - X)^2 / (n - 1)$$

unde
$\Sigma(\chi - X)^2$ - Suma pătratelor abaterilor standard.
(n - 1) - Volumul eşantionului.

## Forma Distribuţiei

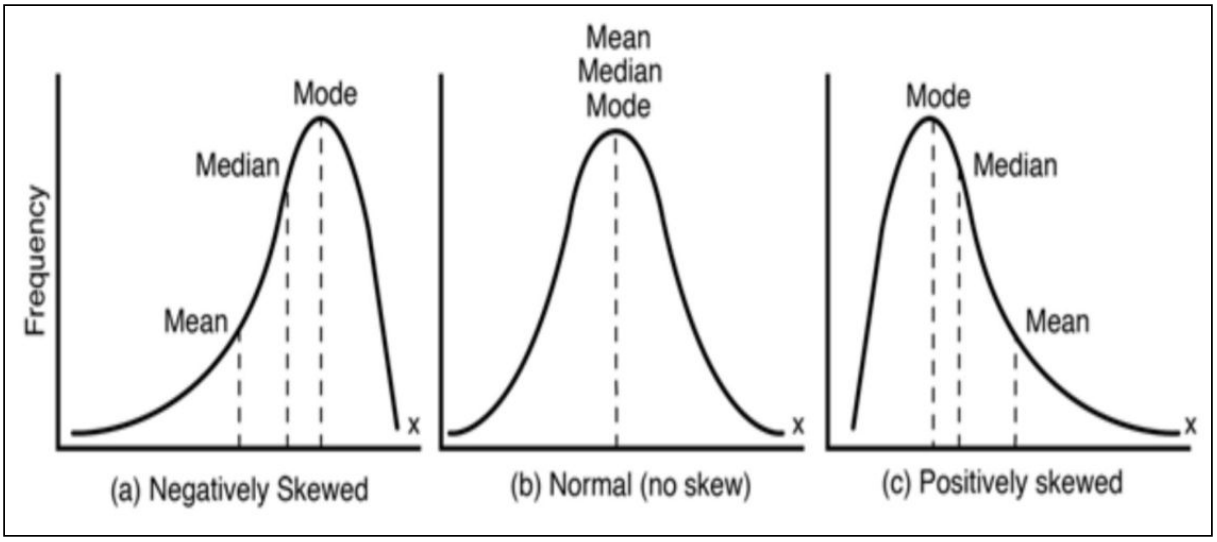
În discuţia privind forma unei distribuţii, vor fi abordate conceptele privind cei doi indicatori ai formei distribuţiei, respectiv modalităţi grafice de reprezentare a unei distribuţii, astfel:

- *Skewness*.
- *Kurtosis*.
- Histograme (*Histograms*).
- Reprezentări grafice de tip *Stem-and-Leaf*.

### *Skewness*

*Skewness*-ul poate fi înţeles în termeni de grad de asimetrie a unei distribuţii. În acest sens, un *skewness* pozitiv poate indica faptul că distribuţia este plasată majoritar înspre stânga (i.e., conţine valori mari). În acest caz, valoarea mediei este mai mare decât valoarea medianei (i.e., media depinde de fiecare scor al distribuţiei şi este afectată de schimbările semnificative, atunci când eşantionul este mic şi conţine valori extreme). Un *skewness* negativ poate indica faptul că distribuţia este plasată majoritar înspre dreapta (i.e., valori mici). În acest caz, are loc acumularea de valori în partea superioară a distribuţiei, cu mai puţine scoruri în partea inferioară a scalei (Sykes et al., 2016).

Pentru reprezentarea conceptului de *skewness*, se va face apel la figura lui Durkhure şi Lodwal (2014, în Junior, Oliviera, & de Faria, 2015, p. 12), ilustrată în continuare.

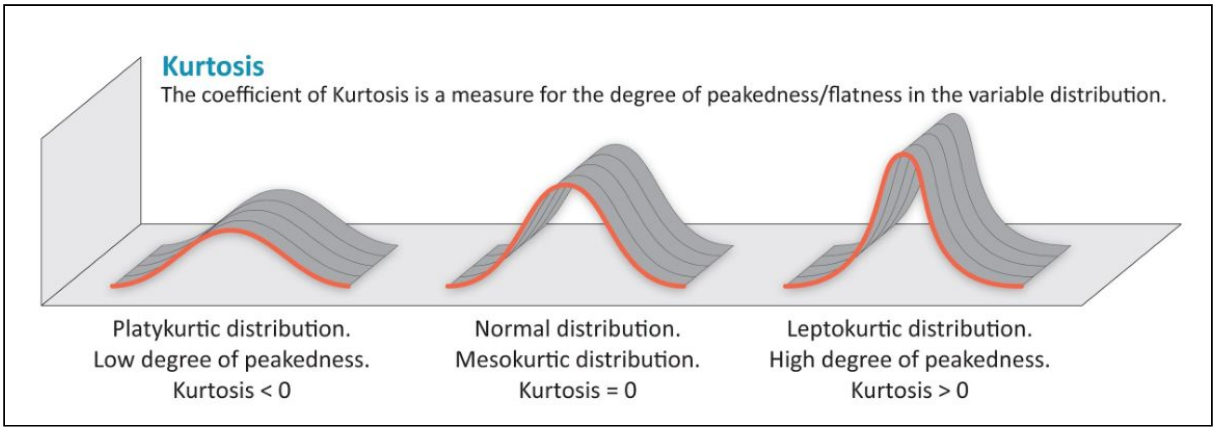



## *Kurtosis*

Conceptul de *kurtosis* se referă la cum sunt concentrate scorurile, în cadrul distribuţiei, la limitele superioare şi inferioare, respectiv la distanţele dintre centru şi limite (denumite "*shoulders*") (Sykes et al., 2016).

Literatura de specialitate utilizează termenul "*peakedness*", asumând faptul că indicatorul *kurtosis* reprezintă cât de "ascuţită" poate fi o distribuţie. Cu toate acestea, indicatorul *kurtosis* nu spune ceva anume despre forma vârfului distribuţiei. În ciuda acestui fapt, este acceptată, încă, interpretarea ambiguă a indicatorului *kurtosis* drept *peakedness* (pentru mai multe detalii, vezi Westfall, 2014).

În fine, pentru înţelegerea conceptului de *kurtosis*, se va face apel la următoarea figură (Kurtosis, n.d.).

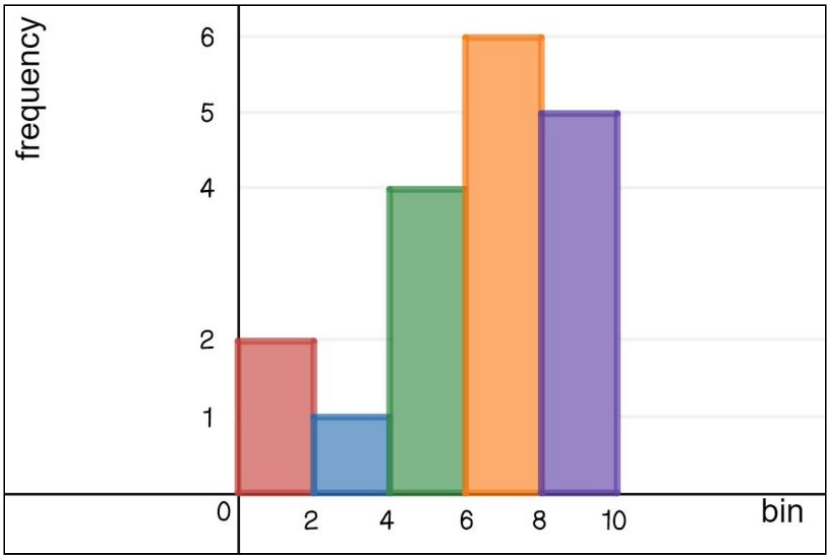



## Histograme (*Histograms*)

Histogramele pot fi înţelese în termeni de reprezentări grafice ale informaţiilor despre frecvenţele anumitor distribuţii sau despre frecvenţele relative ale distribuţiilor unor seturi de date

(Tu, 2008). Astfel, histogramele reprezintă metode grafice ce utilizează bare pentru a prezenta numărul sau frecvenţa datelor. Variabila independentă constă în date de tip interval-raport, reprezentate pe axa absciselor (i.e., axa X), respectiv date de frecvenţă, reprezentate pe axa ordonatelor (i.e., axa Y) (Salkind, 2010d).

Pentru înţelegerea conceptului de histograme, se va face apel la figura propusă de Federspiel (n.d.), după cum urmează.




## Reprezentări Grafice de Tip *Stem-and-Leaf*

Acestea sunt modalităţi grafice descoperite de John Tukey, utilizate pentru analiza exploratorie a datelor continue (Harrington, 2008). Reprezentările grafice de tip *stem-and-leaf* pot fi înţelese ca tabele în care datele sunt împărţite în categoriile "*stem*" sau "*leaf*". În acest sens, valorile din categoria "*stem*" sunt situate pe axa verticală, în timp ce valorile din categoria "*leaf*" sunt situate pe axa orizontală (stem & leaf plot, n.d.)

Pentru înţelegerea conceptului de grafic *stem-and-leaf*, se va face apel la figurile propuse de Meery (2016), după cum urmează.

| Stem | Leaf |
|---|---|
| 0 | 9 |
| 1 | 2, 8, 3, 9, 5 |
| 2 | 0, 3, 9, 2 |
| 3 | 9, 8, 2, 5, 7, 8, 8 |
| 4 | 9, 3, 2, 6 |
| 5 | 8, 9, 0, 3, 5 |
| 6 | 6, 7, 5 |
| 7 | 1 |

| Stem | Leaf |
|---|---|
| 0 | 9 |
| 1 | 2, 3, 5, 8, 9 |
| 2 | 0, 2, 3, 9 |
| 3 | 2, 5, 7, 8, 8, 8, 9 |
| 4 | 2, 3, 6, 9 |
| 5 | 0, 3, 5, 8, 9 |
| 6 | 5, 6, 7 |
| 7 | 1 |

## Bibliografie Selectivă - Statistici Descriptive Univariate


Constantin, D., & Tresset, G. (n.d.). *Practical data analysis* [PDF]. https://www.equipes.lps.u-psud.fr/constantin/IMG/pdf/slides.pdf

Federspiel, A. (n.d.). *What Is a Histogram?*. Retrieved from https://www.expii.com/t/what-is-a-histogram-4635

Galarnyk, M. (2018, September 12). *Understanding Boxplots*. Towards data science. Retrieved from https://towardsdatascience.com/understanding-boxplots-5e2df7bcbd51

Junior, S. P., Oliveira, F. B., & de Faria, M. T. C. (2015). Acoustic Emission Tests on the Analysis of Cracked Shafts of Different Crack Depths. In *23rd ABCM International Congress of Mechanical Engineering* (Vol. 6, pp. 11-18). https://doi.org/10.20906/CPS/COB-2015-1434

Kurtosis (n.d.). O'Reilly. Retrieved from https://www.oreilly.com/library/view/machine-learning-for/9781786469878/f2cbeb0b-4094-4cb9-be54-2d046ff79d7f.xhtml

Meery, B. (2016, January 09). *Stem-and-Leaf Plots*. Retrieved from https://www.ck12.org/c/statistics/understand-and-create-stem-and-leaf-plots/lesson/Stem-and-Leaf-Plots-BSC-PST/

Singh, N., & Oorkavalan, U. (2018). Triple Threshold Statistical Detection filter for removing high density random-valued impulse noise in images. *EURASIP Journal on Image and Video Processing*, *2018*(22), 1-16. https://doi.org/10.1186/s13640-018-0263-0

Standard Deviation Calculator (n.d.). *Standard Deviation Formula*. Retrieved from https://standarddeviationcalc.com/


## Resurse Online - Statistici Descriptive Univariate

1. Descriptive Statistics | Research Methods Knowledge Base
2. Understanding Descriptive and Inferential Statistics
3. Descriptive Statistics
4. Descriptive Statistics
5. Types of Descriptive Statistics – Baseline Help Center
6. Descriptive Statistics: Definition & Charts and Graphs
7. Descriptive Statistics and Graphical Displays | Circulation
8. Understanding Descriptive Statistics | by Sarang Narkhede

# Teste Statistice Elementare pentru Psihologie

## Testul *t* pentru Un Singur Eşantion

Testul *t* pentru un singur eşantion presupune compararea mediei eşantionului cu o valoare specifică (i.e., media populaţiei), cunoscută *a priori* (i.e., înainte de derularea propriu-zisă a testului). În acest sens, un cercetător poate fi interesat dacă media eşantionului diferă semnificativ de media populaţiei (Gerald, 2018).

Asumpţiile testului *t* pentru un singur eşantion, conform lui Gerald (2018) sunt:

- Variabila dependentă este distribuită normal de-a lungul populaţiei.
- Datele colectate sunt independente (i.e., variantele de răspuns ale participanţilor sunt independente).

Important de menţionat este că, înaintea derulării testului statistic în sine, este necesară testarea normalităţii distribuţiei de scoruri (Gerald, 2018).

## Testul *t* pentru Eşantioane Perechi

### Asumpţiile Testului

Testul *t* pentru eşantioane perechi poate fi întâlnit în literatura de specialitate sub denumirile de "*dependent t-test*" şi "*paired sample t-test*". Acest test constă în compararea a două grupuri de scoruri şi a mediilor acestora pentru a observa diferenţe între grupuri (Gerald, 2018). Este important de cunoscut faptul că datele colectate nu sunt independente (i.e., respondenţii dintr-un grup au o oarecare legătură cu cei din alt grup) (University of Arizona Military Reach, 2009, în Gerald, 2018). Deseori, testul *t* pentru eşantioane perechi este utilizat în design-uri pretest-posttest (i.e., măsurători ale aceloraşi constructe pe aceiaşi participanţi, înainte şi după o intervenţie), unde scorurile sunt interdependente (Gerald, 2018).

Marilyn şi Theresa (2003, în Gerald, 2018) delimitează o serie de proceduri necesare în derularea unui test *t* pentru eşantioane perechi, după cum urmează:

- Enunţarea ipotezei de nul ($H_0$) şi a ipotezei cercetării (i.e., denumită şi "ipoteză alternativă") ($H_1$). Ipoteza alternativă este opusul ipotezei de nul. Exemplificare:
  - $H_0$: Nu există diferenţe semnificative între mediile celor două eşantioane.
  - $H_1$: Există diferenţe semnificative între mediile celor două eşantioane.
- Specificarea nivelului de semnificaţie statistică (e.g., $\alpha = .05$, bilateral). Acesta reprezintă probabilitatea maximă tolerată pentru respingerea ipotezei de nul.
- Calcularea valorii testului (*t*).
- Compararea valorii obţinute a testului cu o valoare tabelară (vezi https://home.ubalt.edu/ntsbarsh/Business-stat/StatistialTables.pdf, p. 2).
- Interpretarea rezultatului obţinut.

### Mărimea Efectului

Cea mai des utilizată modalitate de calculare a mărimii efectului testului *t* pentru eşantioane perechi este *Eta Squared* (Pallant, 2007, în Gerald, 2018). Pentru formula lui *Eta Squared*, alături de indicaţiile aferente de interpretare a valorii acestei metrici, Gerald (2018, p. 53) propune următoarele.

$$Eta\ Squared = \frac{1}{1 + \frac{n-1}{t^2}}$$

Where
t is the calculated value of the dependent sample t-test
n is the sample size
The calculated value of the Eta Squared should be between 0 and 1.
0.01 Small effect
0.06 Moderate effect
0.14 Large effect

## Testul *t* pentru Eşantioane Independente

### Asumpţiile Testului

Testul *t* pentru eşantioane independente presupune compararea a două grupuri independente ale căror medii sunt independente una de cealaltă (University of Arizona Military Reach, 2009, în Gerald, 2018). În acest sens, grupurile nu pot fi asociate în vreun fel (e.g., perechi), în cadrul acestui test. Procedurile necesare pentru derularea acestui test sunt identice cu cele ale testului *t* pentru eşantioane perechi (Gerald, 2018).

### Mărimea Efectului

Andy (2009, în Gerald, 2018) defineşte mărimea efectului ca fiind o măsură obiectivă şi, deseori, standardizată, a mărimii sau magnitudinii unui efect observat. Cele mai des întâlnite metrici, în acest sens, sunt: (1) *d*-ul lui Cohen; (2) coeficientul *r* de corelaţie Pearson; şi (3) *Odds Ratio* (*OR*) (Gerald, 2018).

Gerald (2018, p. 53) propune următoarea formulă, însoţită de indicaţii de interpretare a coeficientului *r*.

$$r = \sqrt{\frac{t^2}{t^2 + n_1 + n_2 - 2}}$$

Where
$t$ is the calculated value of the independent sample t-test
$n$ is the sample size
$r = 0.1$ (Small effect)
$r = 0.3$ (Medium effect)
$r = 0.5$ (Large effect)

# Analiza de Varianţă (*ANOVA*) Unifactorială

## Asumpţiile Testului

Analiza de varianţă unifactorială (*ANOVA One-Way*) reprezintă o tehnică de generalizare a asumpţiilor testului *t* de la două eşantioane la trei sau mai multe (Heiberger & Neuwirth, 2009).

Ipotezele testului *ANOVA One-Way* pot fi formulate după cum urmează:

- $H_0$: Nu există diferenţe semnificative între mediile eşantioanelor analizate.
- $H_1$: Există diferenţe semnificative între mediile eşantioanelor analizate.

Conform lui Kim (2017), asumpţiile testului *ANOVA One-Way* sunt:

- Normalitate (i.e., evaluată fie cu ajutorul testului Shapiro-Wilk, fie cu testul Kolmogorov, fie vizual, cu ajutorul graficelor de tip *Q-Q*).
- Independenţa grupurilor analizate.
- Egalitatea (i.e., denumită şi "omogenitate") varianţelor eşantioanelor (i.e., evaluată cu ajutorul testului lui Levene).

## Mărimea Efectului

Cele mai des utilizate metrici, conform lui Okada (2013), de stabilire a mărimii efectului pentru testul *ANOVA* sunt:

- *Epsilon Squared* ($\varepsilon^2$).
- *Eta Squared* ($\eta^2$).
- *Omega Squared* ($\omega^2$).

### *Epsilon Squared*

Acest indicator conţine cea mai mică eroare dintre toate cele trei specificate (Okada, 2013).

### *Eta Squared*

Acest indicator are cea mai mică abatere standard dintre toate cele trei specificate. Cu toate acestea, nu reprezintă, neapărat, un bun estimator (i.e., poate fi observată o cantitate considerabilă de eroare atunci când dimensiunea eşantionului este redusă) (Okada, 2013).

### *Omega Squared*

Acest indicator are cel mai mic *RMSE* (engl., *Root Mean Squared Errors*), pentru toate condiţiile (Okada, 2013).

# Corelaţie

## Asumpţiile Testului

Corelaţia poate fi înţeleasă în termeni de analiză statistică bivariată ce indică magnitudinea asocierii dintre două sau mai multe variabile. În acest sens, o corelaţie pozitivă indică măsura în care scorurile variabilelor cresc sau descresc concomitent, în paralel, în timp ce o corelaţie negativă indică

măsura în care scorurile constructelor variază invers proporţional (i.e., în timp ce valoarea unuia descreşte, valoarea celuilalt creşte şi reciproc valabil) (Zaid, 2015).

Un aspect important de luat în calcul este acela conform căruia corelaţia nu înseamna sau nu presupune cauzalitate. Adiţional, valorile coeficienţilor de corelaţie pot varia între -1 (i.e., corelaţie negativă perfectă) şi +1 (i.e., corelaţie pozitivă perfectă). Valoarea 0 indică faptul că nu există o corelaţie între variabile (Zaid, 2015).

Conform lui Zaid (2015), asumpţiile testului de corelaţie sunt:

- Variabilele analizate sunt independente.
- Variabilele sunt selectate aleatoriu dintr-o populaţie.
- Omogenitatea (i.e., cunoscută şi sub denumirea de "homoscedasticitate") varianţei datelor (i.e., datele prezintă aceeaşi abatere standard, în grupuri diferite).
- Inexistenţa sau existenţa limitată a cazurilor extreme (engl., "*outliers*").

Coeficienţii testului de corelaţie, conform lui Zaid (2015), sunt:

- Coeficientul *r* al lui Pearson.
- Coeficientul $\tau$ (*tau*) al lui Kendall.
- Coeficientul $\rho$ (*rho*) al lui Spearman.

Formulele coeficienţilor de corelaţie anterior menţionaţi sunt ilustraţi, după cum urmează (Zaid, 2015, pp. 5-11).

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

$$\tau = \frac{n_c - n_d}{\frac{1}{2}n(n-1)}$$

Where:

$\tau$ = Kendall rank correlation coefficient

$n_c$ = number of concordant (Ordered in the same way).

$n_d$ = Number of discordant (Ordered differently).

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

Where:

$\rho$ = Spearman rank correlation coefficient

di= the difference between the ranks of corresponding values Xi and Yi

n= number of value in each data set.

# Regresie Liniară Simplă

## Asumpţiile Testului

Analiza de regresie liniară simplă presupune identificarea şi evaluarea unei relaţii dintre una sau mai multe variabile independente (i.e., predictori, antecedente) şi una dependentă (i.e., criteriu, consecinţă). Aşadar, analiza de regresie liniară simplă presupune testarea cauzalităţii dintre variabilele independente (i.e., caracteristici ce pot fi măsurate direct) şi cele dependente (i.e., caracteristici ale căror valori depind de valorile variabilelor independente) (Zaid, 2015).

Analiza de regresie liniară simplă operează în baza unor întrebări ce ţin de fidelitate şi de validitate, concepte învăţate şi/ sau repetate în capitolele anterioare. În acest sens, Zaid (2015) propune următoarele întrebări, pentru o mai bună înţelegere a analizei de regresie liniară simplă:

- Modelul testat are sens intuitiv? Este modelul uşor de înţeles şi de interpretat?
- Sunt toţi coeficienţii semnificativi statistic (i.e., valori $p < .05$)?
- Au sens semnele (+/ -) asociate cu coeficienţii (i.e., pentru predicţii pozitive/ negative)?
- Modelul prezice valori apropiate de realitate (i.e., semnificaţie practică/ clinică)?
- Este modelul suficient de precis (e.g., indice $R^2$ crescut, eroare standard scăzută, etc.)?

Conform întrebărilor anterioare, Zaid (2015) propune obiectivele analizei de regresie liniară simplă, după cum urmează:

- Explicarea variabilităţii variabilei dependente cu ajutorul variabilei independente.
- Analizarea asocierilor dintre variabile cu scopul de a soluţiona cât de mult variază (i.e., cât de mult "se schimbă") variabila dependentă în raport cu cea independentă.
- Prezicerea variabilei dependente în baza valorilor variabilei independente.

Astfel, asumpţiile analizei de regresie liniară simplă sunt (Zaid, 2015):

- Relaţia dintre variabila independentă şi ce dependentă este una liniară.
- Valoarea aşteptată a erorii de predicţie (i.e., zgomot, *error term*) este 0.
- Varianţa erorii de predicţie este constantă (i.e., asumpţia homoscedasticităţii).
- Nu există autocorelaţie (i.e., corelaţia dintre fix aceeaşi variabilă şi ea însăşi, în momente diferite de timp).
- Variabila independentă nu este corelată cu eroarea de predicţie.
- Eroarea de predicţie este distribuită normal.
- În medie, diferenţa dintre valoarea observată şi valoarea prezisă este 0 (i.e., predicţia făcută în baza analizei de regresie liniară simplă coincide cu realitatea).
- În medie, valorile estimate ale erorilor şi valorile variabilelor independente nu sunt asociate (i.e., predicţia făcută este un efect real, nu zgomot sau eroare).
- Diferenţele pătratelor valorii observate şi valorii prezise sunt similare (i.e., modelul de predicţie se "mulează" bine pe realitate).
- Există o oarecare variaţie a variabilei independente. Dacă în ecuaţie apar mai mult de două variabile, acestea nu ar trebui să fie corelate perfect (i.e., ca o variabilă să o prezică pe alta 100% ar însemna că ele coincid - nu ar mai fi un model de predicţie).

Pentru partea practică a analizei de regresie liniară simplă, este nevoie de cunoaşterea a trei concepte de bază: (1) Constanta (i.e., cunoscută drept "*Intercept*" sau "*Constant*"); (2) Panta (i.e.,

cunoscută drept "*Slope*"); și (3) ecuația dreptei de regresie. Acestea vor fi detaliate, conform lui Zaid (2015), în continuare.

## Constanta (*Intercept/ Constant*)

- Punctul în care regresia se intersectează cu axa Y.
- Măsură a mediei variabilei dependente atunci când panta sau pantele sunt 0.
- Dacă panta sau pantele nu sunt 0, atunci:

*Intercept = Media variabilei dependente - Slope * Media variabilei independente*

## Panta (*Slope*)

- O schimbare a variabilei dependente, concomitentă cu o schimbare a variabilei independente.
- În cazul în care panta este 0, acest fapt sugerează că variabila independentă nu are influență asupra variabilei dependente

## Ecuația Dreptei de Regresie (Zaid, 2015, pp. 14-15)

- One variable, denoted ($x$), is regarded as the **predictor**, **explanatory**, or **independent** variable.
- The other variable, denoted ($y$), is regarded as the **response**, **outcome**, or **dependent** variable.

Mathematically, the regression model is represented by the following equation:

$$y = \beta_0 \pm \beta_1 x_1 \pm \varepsilon_1$$

***Where***

- $x$ independent variable.
- $y$ dependent variable.
- $\beta_1$ The Slope of the regression line
- $\beta_0$ The intercept point of the regression line and the y axis.
- $n$ Number of cases or individuals.
- $\sum xy$ Sum of the product of dependent and independent variables.
- $\sum x$ = Sum of independent variable.
- $\sum y$ = Sum of dependent variable.
- $\sum x^2$ = Sum of square independent variable.

$$\beta_1 = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

## Mărimea Efectului

Indicele de mărime a efectului asociat cu analiza de regresie liniară simplă este coeficientul de determinare ($R^2$). Pentru calcularea acestuia şi pentru calcularea şi explicarea valorii sale ajustate (*Adjusted* $R^2$), se va face apel la consemnările lui Zaid (2015, pp. 17-20), după cum urmează.

$$R^2 = \frac{Explained\ Variation}{Total\ Variation} = \frac{Regression\ Sum\ of\ Square\ (SSR)}{Total\ Sum\ of\ Square\ (SST)}$$

The adjusted R Square value is adjusted for the number of variables included in the regression equation. This is used to estimate the expected shrinkage in R Square that would not generalize to the population because our solution is over-fitted to the data set by including too many independent variables. If the adjusted R Square value is much lower than the R Square value, it is an indication that our regression equation may be over-fitted to the sample, and of limited generalize ability.

$$AdjR^2 = 1 - \frac{n-1}{n-k} * (1 - R^2)$$

## Sumar - Teste Statistice Elementare pentru Psihologie

- Testul *t* pentru un singur eşantion are două asumpţii: (1) compararea mediei eşantionului cu media populaţiei; şi (2) normalitate.
- Testul *t* pentru eşantioane perechi presupune compararea scorurilor sau a mediilor a două grupuri interconectate. În cadrul acestui test, sunt importante următoarele două metrici: (1) valoarea testului (*t*); şi (2) *Eta Squared*.
- Testul *t* pentru eşantioane independente presupune compararea scorurilor sau a mediilor a două grupuri independente. Cele mai des utilizate metrici, în acest sens, sunt: *d*-ul lui Cohen, coeficientul *r* de corelaţie al lui Pearson şi *Odds Ratio* (*OR*).
- Analiza de varianţă unifactorială presupune compararea scorurilor sau a mediilor a trei sau mai multe eşantioane. În acest sens, cele mai des utilizate teste de normalitate sunt *Shapiro-Wilk* şi *Kolmogorov*, cel mai des utilizat test pentru egalitatea varianţelor eşantioanelor este testul *Levene* şi cele mai des întâlnite metrici ale mărimii efectului sunt *Epsilon Squared*, *Eta Squared* şi *Omega Squared*.
- Corelaţia reprezintă o analiză statistică bivariată.
- Corelaţia nu înseamnă cauzalitate.
- $-1 < r/\ \tau/\ \rho < 1$.
- Regresia liniară simplă: Predictori/ Antecedente (*X*) → Criterii/ Consecinţe (*Y*).
- În regresia liniară simplă: $Y = Intercept \pm Slope^{*}X \pm Eroare\ de\ predicţie$.

## Resurse Online - Teste Statistice Elementare pentru Psihologie

1. https://www.minitab.com.au/uploadedFiles/Documents/sample-materials/MortgageProcessTime-EN.pdf
2. One-Sample T-Test in: Basic and Advanced Statistical Tests
3. STATISTICAL TABLES
4. Dependent T-Test - An introduction to when to use this test and what are the variables required
5. T-Test Calculator for 2 Dependent Means
6. Paired Sample T-Test
7. Independent T-Test - An introduction to when to use this test and what are the variables required
8. Independent Samples T Test (Unpaired Samples): Definition, Running
9. The Independent T-test The t-test assesses whether the means of two groups, or conditions, are statistically different from one
10. https://psygnificant.files.wordpress.com/2013/01/st1_09_anova.pdf
11. One-way ANOVA - An introduction to when you should run this test and the test hypothesis
12. One-way ANOVA | When and How to Use It (With Examples)
13. 11. Correlation and regression
14. What is Correlation? | Correlation Examples
15. Correlation | Research Methods Knowledge Base
16. Correlation Game
17. Statistics 202
18. Simple linear regression
19. Simple linear regression

# Analiza de Moderare și Analiza de Mediere

## Analiza de Moderare

### Definiţie şi Caracteristici

Moderarea are loc atunci când relaţia dintre o variabilă independentă (i.e., predictor, antecedent) şi o variabilă dependentă (i.e., criteriu, consecinţă) se modifică, în funcţie de valoarea unei variabile moderatoare (Dawson, 2014, în Memon et al., 2019).

Variabila moderatoare poate fi înţeleasă în termenii unei variabile ce "influenţează natura (e.g., magnitudine şi/ sau direcţia) efectului dintre antecedent şi consecinţă" (Anguinis, Edwards, & Bradley, 2017, p. 2, în Memon et al., 2019, pp. 1-2).

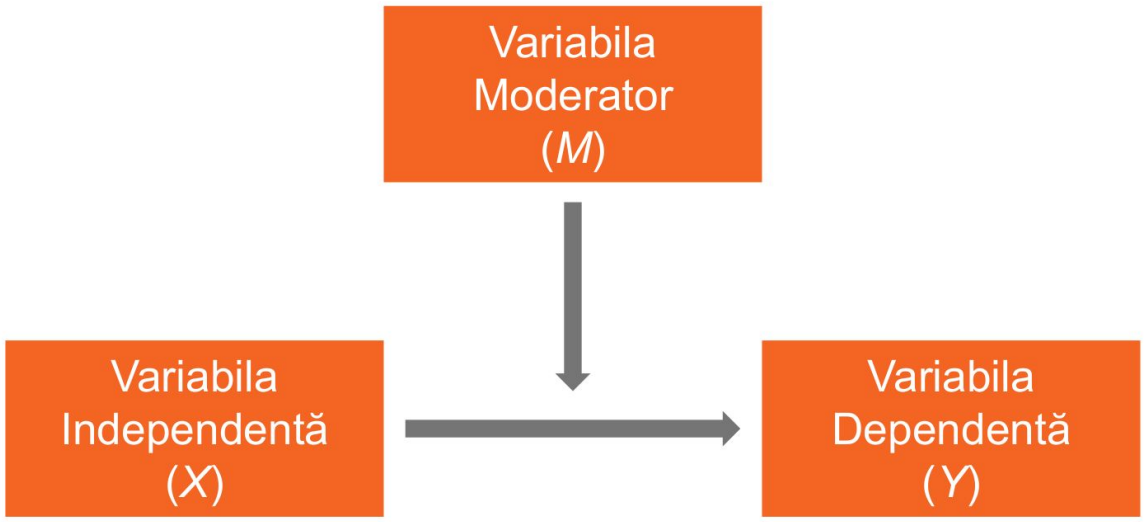
### Asumpţii

Warner (2013) listează o serie de asumpţii ale analizei de moderare, după cum urmează:

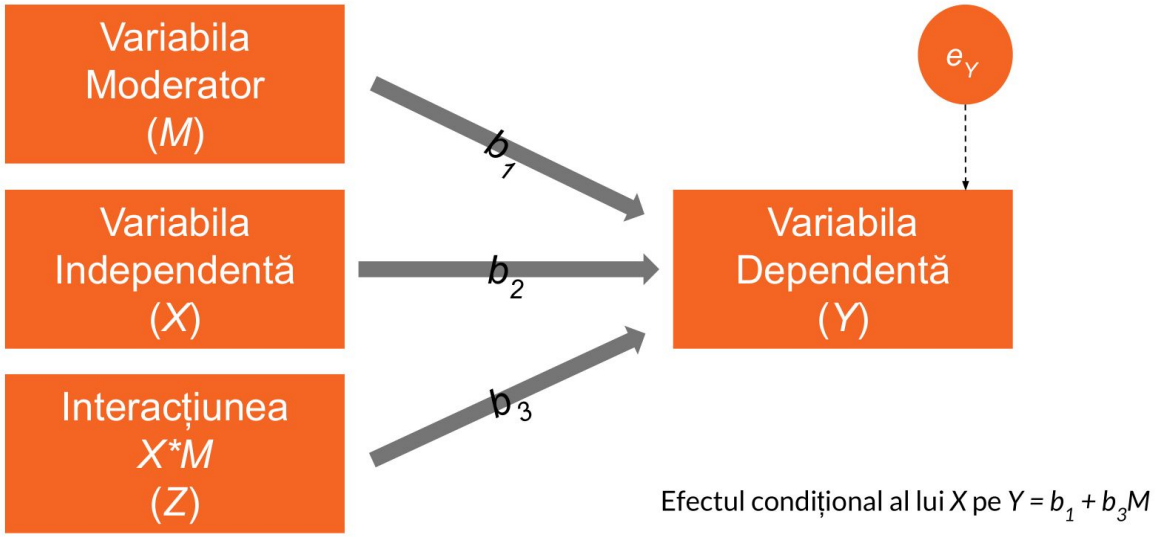
- Scorurile pe toate variabilele cantitative ar trebui să fie distribuite aproximativ normal.
- Asocierile dintre toate perechile de variabile ar trebui să fie liniare.
- Nu ar trebui să existe cazuri univariate sau multivariate extreme (*outliers*).
- Pentru a putea fi utilizată regresia liniară *OLS* (*Ordinary Least Squares*), variabila dependentă (*Y*) trebuie să fie cantitativă.
- Dacă variabila dependentă (*Y*) este categorială, vor fi utilizate metode de regresie logistică.
- Pentru variabilele independente (*X*) categoriale, numărul de scoruri din fiecare grup ar trebui să fie îndeajuns de mare pentru a putea fi obţinute limite relativ restrânse ale intervalelor de încredere.
- Omogenitatea pantelor de regresie, pentru modelele de regresie multiplă (i.e., valoarea parţială a pantei variabilei independente $X_1$ este aceeaşi pe parcursul tuturor valorilor variabilei independente $X_2$).

Pentru o înţelegere aplicată a conceptului de analiză de moderare, vor fi ilustrate, în continuare, modelul conceptual şi modelul statistic aferente acesteia. Totodată, lecturarea articolului lui Hochwater, Witt şi Kacmar (2000) poate ajuta la aprofundarea conceptului.

Model Conceptual ("*Model 1*", Hayes, 2017, în https://osf.io/6kbgj/, p. 6)




Model Statistic ("*Model 1*", Hayes, 2017, în https://osf.io/6kbgj/, p. 7)




# Analiza de Mediere

## Definiţie şi Caracteristici

Analiza de mediere poate fi înţeleasă prin prisma faptului că o variabilă mediatoare ($M_i$) transmite efectul variabilei independente ($X$) pe variabila dependentă ($Y$) (i.e., efectul dintre $X$ şi $Y$ se face prin $M_i$). În forma sa cea mai simplă, medierea reprezintă adiţia unei a treia variabile în relaţia $X$

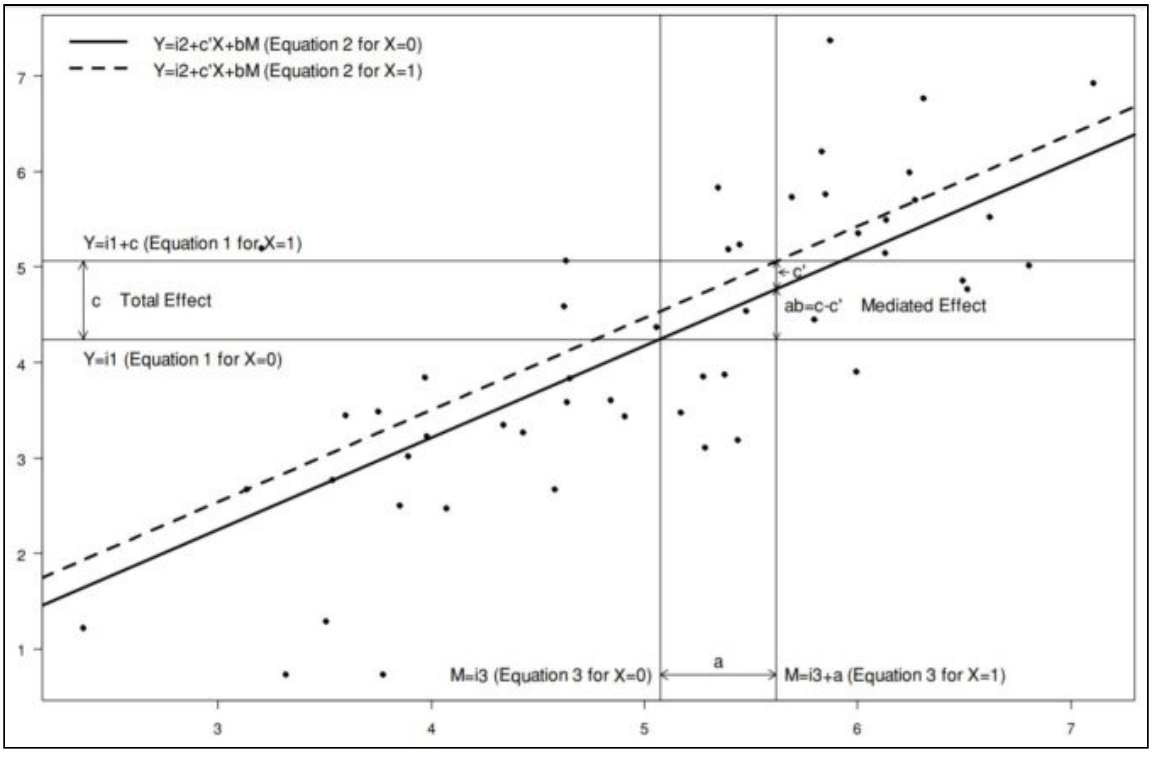
→ $Y$ (i.e., $X \rightarrow M_i \rightarrow Y$, $X$ îl prezice pe $M_i$, $M_i$ îl prezice pe $Y$). Diferenţa dintre variabila mediatoare ($M_i$) şi variabila moderatoare ($M$) este aceea că $M_i$ se află într-o secvenţă cauzală, în timp ce $M$ nu face parte dintr-o astfel de secvenţă (MacKinnon, Fairchild, & Fritz, 2007).

## Asumpţii

MacKinnon şi colaboratorii săi (2007) delimitează o serie de asumpţii ale analizei de mediere, după cum urmează:

- Ecuaţiile de regresie ale medierii:
  - $Y = i_1 + cX + e_1$ (1)
  - $Y = i_2 + c'X + bM_i + e_2$ (2)
  - $M_i = i_3 + aX + e_3$ (3)
- Cantităţile reziduale ($e$-urile) din ecuaţiile (2) şi (3) sunt independente şi $M_i$ şi $e_2$ sunt independente (McDonald, 1997; Merrill, unpublished dissertation, în MacKinnon et al., 2007).
- Nu există o interacţiune $XM_i$ în ecuaţia (3).
- Nu există erori de specificare a ordinii cauzale (e.g., $Y \rightarrow M_i \rightarrow X$, în loc de $X \rightarrow M_i \rightarrow Y$).
- Nu există erori de specificare a direcţiei cauzale (e.g., asociere cauzală reciprocă între $M_i$ şi $Y$).
- Nu există erori de specificare cauzate de variabile absente din măsurare sau de măsurare imperfectă (Holland, 1988; James & Brett, 1984; McDonald, 1997, în MacKinnon et al., 2007).

Pentru vizualizarea acestor asumpţii, se va face apel la figura graficului efectului de mediere, propusă de MacKinnon şi colaboratorii acestuia (2007, p. 600), după cum urmează.

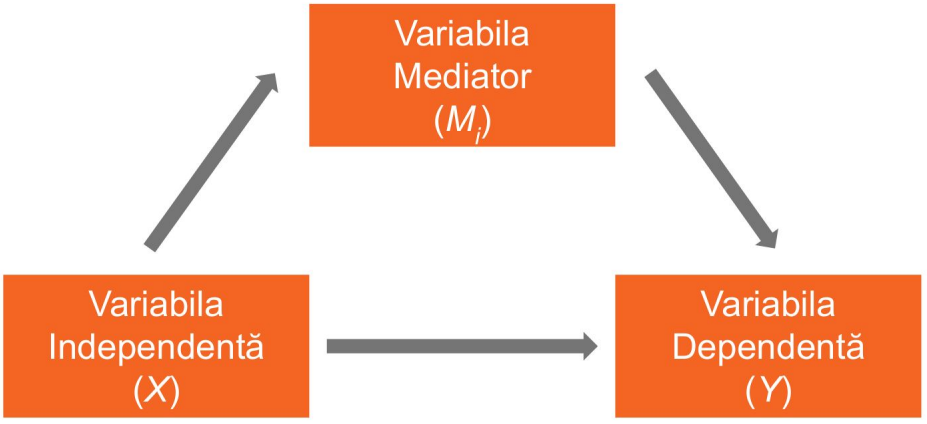


## Mediere Totală vs. Mediere Parţială

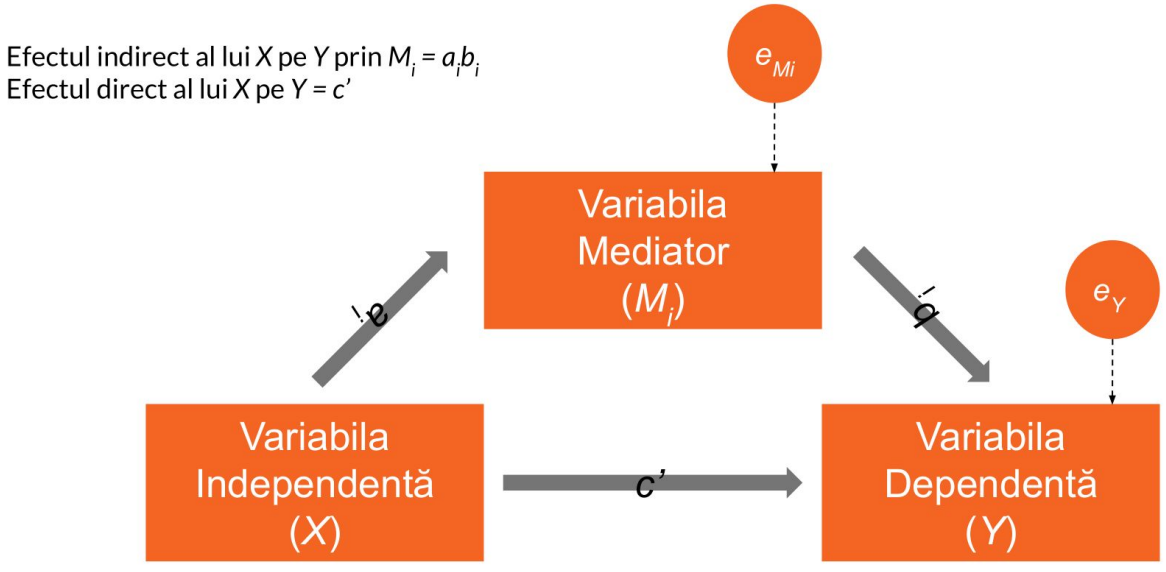
În cazul medierii totale, coeficientul *c'* nu este semnificativ statistic. În contrast, în cazul medierii parţiale, coeficientul *c'* este semnificativ statistic (MacKinnon et al., 2007).

Pentru vizualizarea interacţiunilor dintre variabile, se va face apel la modelul conceptual şi la modelul statistic al analizei de mediere, fapt urmat de recomandarea unui articol (Singh & Singh, 2018) din literatura de specialitate, pentru observarea aplicabilităţii acestei analize statistice.

Model Conceptual ("*Model 4*", Hayes, 2017, în https://osf.io/6kbgj/, p. 15)




Model Statistic ("*Model 4*", Hayes, 2017, în https://osf.io/6kbgj/, p. 16)



## Bibliografie Selectivă - Analiza de Moderare şi Analiza de Mediere

Hochwarter, W. A., Witt, L. A., & Kacmar, K. M. (2000). Perceptions of organizational politics as a moderator of the relationship between consciousness and job performance. *Journal of Applied Psychology*, *85*(3), 472–478. https://doi.org/10.1037/0021-9010.85.3.472

Singh, V., & Singh, M. (2018). A burnout model of job crafting: Multiple mediator effects on job performance. *IIMB Management Review*, *30*(4), 305-315. https://doi.org/10.1016/j.iimb.2018.05.001

## Resurse Online - Analiza de Moderare şi Analiza de Mediere

1. Moderation and Mediation
2. What is the difference between moderation and mediation?
3. Chapter 14: Mediation and Moderation
4. What are the differences between moderation and mediation effects?
5. 3. Moderation and Mediation Explained
6. Mediation and moderation analyses: exploring the complex pathways between hope and quality of life among patients with schizophrenia
7. A General Model for Testing Mediation and Moderation Effects
8. https://link.springer.com/chapter/10.1007%2F978-94-6209-404-8_13

# Semnificația Statistică

## Definiţie

Informal, semnificaţia statistică (i.e., valoarea *p*) reprezintă probabilitatea ca un sumar statistic al datelor, raportat la un model statistic specific, să fie egal cu sau situat extrem faţă de valoarea sa observată (Wasserstein & Lazar, 2016).

## Accepţia Clasică

În accepţia clasică a conceptului, semnificaţia statistică reprezintă "probabilitatea (riscul) de a respinge, în mod eronat, ipoteza nulă deşi aceasta este adevărată: nivelul de semnificaţie măsoară riscul de apariţie a unei erori de tip I. Un fapt statistic este semnificativ dacă produce respingerea ipotezei nule. Este de notat, totuşi, că un fapt semnificativ statistic nu este, în mod neapărat, semnificativ în lumea reală" (Clocotici, n.d.). Pentru o modalitate clasică de interpretare a nivelului de semnificaţie statistică, este recomandată lecturarea materialului lui Tenny şi Abdelgawad (2020).

## Accepţia Revizuită

### Principiile Semnificaţiei Statistice

Wasserstein şi Lazar (2016) conturează cele 6 principii ale semnificaţiei statistice, după cum urmează:

1. Valorile *p* pot indica măsura în care datele sunt incompatibile cu un anumit model statistic.
2. Valorile *p* nu măsoară probabilitatea ca o ipoteză testată să fie adevărată, sau probabilitatea ca datele să fi fost rezultatul şansei.
3. Concluziile ştiinţifice şi deciziile ce ţin de informarea afacerilor sau politicilor nu ar trebui bazate pe ideea conform căreia valoarea *p* trece sau nu de un anumit prag prestabilit (e.g., pragul de 0.05, prin convenţie).
4. Inferenţa precisă necesită o raportare completă şi transparentă.
5. Valorile *p* nu măsoară mărimea unui efect sau importanţa unui rezultat.
6. În sine, o valoare *p* nu reprezintă o metrică satisfăcătoare de evidenţă, în privinţa unui model sau a unei ipoteze.

### Semnificaţia Statistică, Mărimea Efectului şi Semnificaţia Clinică

Având în vedere faptul că semnificaţia statistică arată probabilitatea asociată ipotezei de nul, aceasta nu determină dacă rezultatul este unul important, semnificativ, individual sau îndeajuns de substanţial pentru a prezenta semnificaţie clinică. Pentru interpretarea semnificaţiei clinice, practicienii descriu mărimea efectului unei anumite intervenţii. Aceste efecte sunt, deseori, raportate cu ajutorul unor măsuri standardizate (e.g., indicele *d* al lui Cohen) (Hunter, Miles, Paddock, & D'Amico, 2013).

## Concluzii

Un test de semnificaţie statistică reprezintă un calcul matematic ce ia în considerare testul statistic, ipoteza de nul şi distribuţia testului statistic, raportată la ipoteza de nul. Astfel, semnificaţia statistică reprezintă doar una dintre componentele analizei unui set de date, recomandându-se suplimentarea cu estimări ale mărimilor efectelor, consideraţii privind mărimea eşantionului şi discuţii privind validitatea deciziei statistice. Este important de reţinut că un rezultat semnificativ statistic nu reprezintă, neapărat, un rezultat important, raportat la un anumit context. Cu toate acestea, nivelul de semnificaţie statistică este o metrică ce trebuie luată în considerare, în contextul cercetărilor dintr-un anumit domeniu (Reid, 2001).

## Bibliografie Selectivă - Semnificaţia Statistică

Clocotici, V. (n.d.). Semnificaţie, Nivel de semnificaţie, Semnificaţie statistică. Dicţionar explicativ de statistică. Retrieved from https://profs.info.uaic.ro/~val/statistica/StatGloss.htm

Tenny, S., & Abdelgawad, I.. Statistical Significance. [Updated 2020 Jul 10]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2020 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK459346/

## Resurse Online - Semnificaţia Statistică

1. ASA statement on P-values and statistical significance: Development and impact
2. Statisticians Outline a 'Post p less than 0.05 Era' in New Scientific Statement
3. https://shinyapps.org/apps/p-checker/
4. https://www.shinyapps.org/apps/p-hacker/
5. https://www.shinyapps.org/apps/vs-mpr/

https://shinyapps.org/showapp.php?app=https://tellmi.psy.lmu.de/felix/PPV&by=Michael%20Zehetleitner%20and%20Felix%20Sch%C3%B6nbrodt&title=When%20does%20a%20significant%20p-value%20indicate%20a%20true%20effect?&shorttitle=When%20does%20a%20significant%20p-value%20indicate%20a%20true%20effect?

# Referințe Bibliografice


Adams, H., Cervantes, P., Jang, J., & Dixon, D. (2014). Standardized assessment. *Practical Resources for the Mental Health Professional*, 501-516. https://doi.org/10.1016/B978-0-12-411603-0.00025-2

Baker, S. G., & Lindeman, K. S. (2001). Rethinking historical controls. *Biostatistics*, *2*(4), 383-396. https://doi.org/10.1093/biostatistics/2.4.383

Bartko, J. J. (1966). The intraclass correlation coefficient as a measure of reliability. *Psychological Reports*, *19*(1), 3-11. https://doi.org/10.2466%2Fpr0.1966.19.1.3

Battaglia, M. (2008). Convenience sampling. In P. J. Lavrakas (Ed.), *Encyclopedia of survey research methods* (pp. 149-149). Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781412963947.n105

Beaton, D. E., Boers, M., & Tugwell, P. (2013). 33 - Assessment of health outcomes. In G. S. Firestein, R. C. Budd, … J. R. O'Dell (Eds.), *Kelley's Textbook of Rheumatology* (Ninth Edition) (pp. 462-475). Elsevier. https://doi.org/10.1016/B978-1-4377-1738-9.00033-5

Boston University Medical Campus. (n.d.). *Case reports and case series*. Boston University Medical Campus and Boston Medical Center. http://www.bumc.bu.edu/irb/submission-requirements/special-submission-requirements/case-reports-and-case-series/

Campbell, D. T., & Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio Books. Available at http://davidpassmore.net/courses/data/_book/Camp_and_Stanley.pdf

Cocoş, B. (2020, September 3). *Cercetare în Psihologie*. https://doi.org/10.17605/OSF.IO/BJ6K8

Collins, L. M. (2007). Research design and methods. In J. E. Birren (Ed.), *Encyclopedia of gerontology* (2nd Edition) (pp. 433-442). Elsevier. https://doi.org/10.1016/B0-12-370870-2/00162-1

Cox, B. (2008). Systematic sampling. In P. J. Lavrakas (Ed.), *Encyclopedia of survey research methods* (pp. 872-872). Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781412963947.n569

Crespo-Facorro, B., Bernardo, M., Argimon, J. M., Arrojo, M., Bravo-Ortiz, M. F., Cabrera-Cifuentes, A., ... & Olivares, J. M. (2017). Effectiveness, efficiency and efficacy in the multidimensional treatment of schizophrenia: Rethinking project. *Revista de Psiquiatría y Salud Mental (English Edition)*, *10*(1), 4-20. https://doi.org/10.1016/j.rpsmen.2016.09.002

El-Gilany, A.-H. (2018a). What is case series?. *Asploro Journal of Biomedical and Clinical Case Reports*, *1*(1), 10-15. Retrieved from https://www.researchgate.net/profile/Abdel-Hady_El-Gilany/publication/327449197_What_is_case_series/links/5b90253e299bf114b7f86ab9/What-is-case-series.pdf

El-Gilany, A.-H. (2018b). What is case report?. *Asploro Journal of Biomedical and Clinical Case Reports*, *1*(1), 7-9. Retrieved from https://asploro.com/wp-content/uploads/2018/08/What-is-case-report.pdf

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191. https://doi.org/10.3758/BF03193146

García-Pérez, M. A. (2012). Statistical conclusion validity: Some common threats and simple remedies. *Frontiers in Psychology*, *3*, 1-11. https://doi.org/10.3389/fpsyg.2012.00325

Gerald, B. (2018). A brief review of independent, dependent and one sample t-test. *International Journal of Applied Mathematics and Theoretical Physics*, *4*(2), 50-54. Available at http://article.ijamtp.org/pdf/10.11648.j.ijamtp.20180402.13.pdf

Glen, S. (2018, June 16). *Average inter-item correlation: Definition, example*. StatisticsHowTo. Retrieved from https://www.statisticshowto.com/average-inter-item-correlation/

Goddard III, R. D., & Villanova, P. (2005). Designing surveys and questionnaires for research. In F. T. L. Leong & J. T. Austin (Eds.), *The psychology research handbook: A guide for graduate students and research assistants* (Edition 2, pp. 114-124). SAGE Publications.

Hargreaves, W. A., Catalano, R. A., Hu, T.-W., & Cuffel, B. (1998). Mental Health Services Research. *Comprehensive Clinical Psychology*, 225–241. https://doi.org/10.1016/b0080-4270(73)00275-3

Hariton, E., & Locascio, J. J. (2018). Randomised controlled trials—The gold standard for effectiveness research. *BJOG: An International Journal of Obstetrics and Gynaecology*, *125*(13), 1716. https://doi.org/10.1111%2F1471-0528.15199

Harrington, R. (2008). Stem-and-leaf plot. In S. Boslaugh (Ed.), *Encyclopedia of epidemiology* (Vol. 1, pp. 1001-1001). Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781412953948.n440

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (Second Edition). Guilford Publications. Available at https://www.researchgate.net/profile/Adasa_Nkrumah/post/How_to_run_a_moderation_analysis_with_3_moderators_using_PROCESS-Macro_AHayes/attachment/5e74fc08cfe4a7809f89b936/AS%3A871213949804544%401584725000616/download/templates.pdf

Heiberger, R. M., & Neuwirth, E. (2009). *One-way ANOVA. In R through excel* (pp. 165-191). Springer, New York, NY. https://doi.org/10.1007/978-1-4419-0052-4_7

Hilger, N., Beauducel, A. (2017). Parallel-Forms Reliability. In: Zeigler-Hill V., Shackelford T. (eds) *Encyclopedia of Personality and Individual Differences*. Springer, Cham. https://doi.org/10.1007/978-3-319-28099-8_1337-1

Himmelfarb Health Sciences Library. (n.d.). *Cohort study*. Himmelfarb Health Sciences Library. https://himmelfarb.gwu.edu/tutorials/studydesign101/cohorts.cfm

Hunter, S. B., Miles, J. N. V., Paddock, S. M., & D'Amico, E. J. (2013). Evaluating treatment efficacy. In P. M. Miller (Ed.), *Interventions for addiction. Comprehensive addictive behaviors and disorders* (Volume 3, pp. 589-597). Academic Press. https://doi.org/10.1016/B978-0-12-398338-1.00061-0

Johansson, B., & Wahlin, Å. (1998). Cognition and geropsychological assessment. *Comprehensive Clinical Psychology*, *7*, 25-53. https://doi.org/10.1016/B0080-4270(73)00068-7

Johnson, E. (2013) Face Validity. In: F. R. Volkmar (Eds.) *Encyclopedia of Autism Spectrum Disorders*. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-1698-3_308

Kaur, P., & Stoltzfus, J. (2017). Type I, II, and III statistical errors: A brief overview. *International Journal of Academic Medicine*, *3*(2), 268-270. https://doi.org/10.4103/IJAM.IJAM_92_17

Khorsan, R., & Crawford, C. (2014). External validity and model validity: a conceptual approach for systematic review methodology. *Evidence-Based Complementary and Alternative Medicine*, *2014*, 1-12. https://doi.org/10.1155/2014/694804

Kim, T. K. (2017). Understanding one-way ANOVA using conceptual figures. *Korean Journal of Anesthesiology*, *70*(1), 22-26. https://doi.org/10.4097%2Fkjae.2017.70.1.22

Kirk, R. E. (2003). Experimental Design. In J. A. Schinka & W. F. Velicer (Eds.), *Wiley Handbook of Psychology, Volume 2, Research Methods in Psychology* (pp. 3-32). Hoboken, NJ: Wiley.

Kyonka, E. G. (2019). Tutorial: Small-N power analysis. *Perspectives on Behavior Science*, *42*(1), 133-152. https://doi.org/10.1007/s40614-018-0167-4

Lange R.T. (2011) Inter-rater Reliability. In J. S. Kreutzer, J. DeLuca & B. Caplan (Eds.), *Encyclopedia of Clinical Neuropsychology*. Springer, New York, NY. https://doi.org/10.1007/978-0-387-79948-3_1203

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, *58*, 593–614. https://doi.org/10.1146/annurev.psych.58.110405.085542

Mathers, N., Fox, N., & Hunn, A. (2007). *Surveys and questionnaires*. Trent RDSU. Available at https://www.rds-yh.nihr.ac.uk/wp-content/uploads/2013/05/12_Surveys_and_Questionnaires_Revision_2009.pdf

Maul, A. (2018). Judgment sampling. In B. Frey (Ed.), *The SAGE encyclopedia of educational research, measurement, and evaluation* (Vol. 1, pp. 914-914). Thousand Oaks,, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781506326139.n367

McConaughy, S. H. (2001). Chapter 10 - The Achenbach system of empirically based assessment. In J. J. W. Andrews, D. H. Saklofske & H. L. Janzen (Eds.), *Handbook of psychoeducational assessment. Ability, achievement and behavior in children* (pp. 289-324). Academic Press. https://doi.org/10.1016/B978-012058570-0/50012-4

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica: Biochemia Medica*, *22*(3), 276-282. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3900052/

Memon, M. A., Cheah, J. H., Ramayah, T., Ting, H., Chuah, F., & Cham, T. H. (2019). Moderation analysis: issues and guidelines. *Journal of Applied Structural Equation Modeling*, *3*(1), 1-9. Available at https://www.researchgate.net/profile/Mumtaz_Memon/publication/331175332_Moderation_Analysis_Issues_and_Guidelines/links/5c6abcea299bf1e3a5b012cd/Moderation-Analysis-Issues-and-Guidelines.pdf

Meng, X. (2013, February). Scalable simple random sampling and stratified sampling. In *International Conference on Machine Learning* (pp. 531-539). Available at http://proceedings.mlr.press/v28/meng13a.pdf

Möller, H. J. (2011). Effectiveness studies: Advantages and disadvantages. *Dialogues in Clinical Neuroscience*, *13*(2), 199-207. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3181999/pdf/DialoguesClinNeurosci-13-199.pdf

Multani, P. K., Saini, N., Kaur, R., & Sharma, P. (2014). Biomarkers for drugs of abuse and neuropsychiatric disorders: models and mechanisms. In *Biomarkers in Toxicology* (pp. 983-1001). Academic Press. https://doi.org/10.1016/B978-0-12-404630-6.00059-2

Murad, M. H., Asi, N., Alsawas, M., & Alahdab, F. (2016). New evidence pyramid. *BMJ Evidence-Based Medicine*, *21*(4), 125-127. http://doi.org/10.1136/ebmed-2016-110401

Naderifar, M. & Goli, H., & Ghaljaei, F. (2017). Snowball sampling: A purposeful method of sampling in qualitative research. *Strides in Development of Medical Education*. In Press. https://doi.org/10.5812/sdme.67670

Nishisato, S., & Torii, Y. (1971). Effects of categorizing continuous normal variables on product-moment correlation. *Japanese Psychological Research*, *13*(1), 45-49. https://doi.org/10.4992/psycholres1954.13.45

Nugroho, A., Warnars, H. L. H. S., Heriyadi, Y., & Tanutama, L. (2019, November). Measure The Level Of Success In Using Google Drive with the Kuder Richardson (KR) Reliability Method. In *2019 International Congress on Applied Information Technology (AIT)* (pp. 1-7). IEEE. https://doi.org/10.1109/AIT49014.2019.9144915

Okada, K. (2013). Is omega squared less biased? A comparison of three major effect size indices in one-way ANOVA. *Behaviormetrika*, *40*(2), 129-147. https://doi.org/10.2333/bhmk.40.129

Organisation for Economic Cooperation and Development. (2002, May 21). *Efficiency (in statistical estimation)*. OECD. Retrieved from https://stats.oecd.org/glossary/detail.asp?ID=3746

Ottenbacher, K. J. (1992). Statistical conclusion validity and type IV errors in rehabilitation research. *Archives of Physical Medicine and Rehabilitation*, *73*(2), 121-125. Available at https://www.archives-pmr.org/article/0003-9993(92)90087-D/pdf

Pandis, N., & Cobourne, M. T. (2013). Clinical trial design for orthodontists. *Journal of Orthodontics*, *40*(2), 93-103. https://doi.org/10.1179%2F1465313313Y.0000000048

Patino, C. M., & Ferreira, J. C. (2018). Internal and external validity: can you apply research study results to your patients?. *Jornal Brasileiro de Pneumologia*, *44*(3), 183-183. https://doi.org/10.1590%2FS1806-37562018000000164

Prince, M. (2011). 9 - Epidemiology. In P. Wright, J. Stern & M. Phelan (Eds.), Core Psychiatry (Third Edition, pp. 115-129). Saunders Ltd. https://doi.org/10.1016/B978-0-7020-3397-1.00009-4

Public Health England. (2020, January 30). Before-and-after study: comparative studies. https://www.gov.uk/guidance/before-and-after-study-comparative-studies

Reid, N. (2001). Significance, tests of. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (pp. 14085-14091). Pergamon. https://doi.org/10.1016/B0-08-043076-7/00506-4

Revicki D. (2014) Internal Consistency Reliability. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research*. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0753-5_1494

Rukmana D. (2014) Quota Sampling. In A. C. Michalos (Ed.), *Encyclopedia of Quality of Life and Well-Being Research*. Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0753-5_2393

Salkind, N. J. (2010a). Random assignment. In *Encyclopedia of research design* (Vol. 1, pp. 1194-1197). Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781412961288.n359

Salkind, N. J. (2010b). *Encyclopedia of research design* (Vols. 1-0). Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781412961288

Salkind, N. J. (2010c). Internal consistency reliability. In *Encyclopedia of research design* (Vol. 1, pp. 616-619). Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781412961288.n191

Salkind, N. J. (2010d). Histogram. In *Encyclopedia of research design* (Vol. 1, pp. 572-573). Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781412961288.n177

Sedgwick, P. (2014a). Cross sectional studies: advantages and disadvantages. *BMJ*, *348*, 1-2. https://doi.org/10.1136/bmj.g2276

Sedgwick, P. (2014b). Before and after study designs. *BMJ*, *349*, 1-2. https://doi.org/10.1136/bmj.g5074

Sedgwick, P. (2015). Multistage sampling. *BMJ*, *351*, 1-2. https://doi.org/10.1136/bmj.h4155

Setia, M. S. (2016). Methodology series module 2: Case-control studies. *Indian Journal of Dermatology*, *61*(2), 146-151. https://doi.org/10.4103%2F0019-5154.177773

Souza, A. C. D., Alexandre, N. M. C., & Guirardello, E. D. B. (2017). Psychometric properties in instruments evaluation of reliability and validity. *Epidemiologia e Serviços de Saúde*, *26*, 649-659. https://doi.org/10.5123/S1679-49742017000300022

stem & leaf plot (n.d.). BusinessDictionary.com. Retrieved from http://www.businessdictionary.com/definition/stem-leaf-plot.html

Sykes, L M, Gani, F, & Vally, Z. (2016). Statistical terms Part 1: The meaning of the MEAN, and other statistical terms commonly used in medical research. *South African Dental Journal*, *71*(6), 274-278. Available at http://www.scielo.org.za/pdf/sadj/v71n6/09.pdf

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53-55. https://doi.org/10.5116%2Fijme.4dfb.8dfd

Thompson, C. B., & Panacek, E. A. (2007). Research study designs: Non-experimental. *Air Medical Journal*, *26*(1), 18-22. https://doi.org/10.1016/j.amj.2006.10.003

Tu, R. (2008). Histogram. In S. Boslaugh (Ed.), *Encyclopedia of epidemiology* (Vol. 1, pp. 496-496). Thousand Oaks, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781412953948.n212

Uman, L. S. (2011). Systematic reviews and meta-analyses. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, *20*(1), 57-59. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3024725/

Uttley, J. (2019). Power analysis, sample size, and assessment of statistical assumptions—Improving the evidential value of lighting research. *Leukos*, *15*(2-3), 143-162. https://doi.org/10.1080/15502724.2018.1533851

Wakar, B. & Poznyak, D. (2018). Cluster sampling. In B. Frey (Ed.), *The SAGE encyclopedia of educational research, measurement, and evaluation* (Vol. 1, pp. 299-299). Thousand Oaks,, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781506326139.n117

Warner, R. M. (2013). Moderation: Tests for interaction in multiple regression. In *Applied Statistics From Bivariate Through Multivariate Techniques* (Second Edition, pp. 611-644). Available at https://uk.sagepub.com/sites/default/files/upm-binaries/47569_ch_15.pdf

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: Context, process, and purpose. *The American Statistician*, *70*(2), 129–133. https://doi.org/10.1080/00031305.2016.1154108

Westfall, P. H. (2014). Kurtosis as peakedness, 1905–2014. RIP. *The American Statistician*, *68*(3), 191-195. https://doi.org/10.1080%2F00031305.2014.917055

Zaid, M. A. (2015). *Correlation and regression analysis*. Retrieved from http://www.oicstatcom.org/file/TEXTBOOK-CORRELATION-AND-REGRESSION-ANALYSIS-EGYPT-EN.pdf

Zhang, Y., Hedo, R., Rivera, A., Rull, R., Richardson, S., & Tu, X. M. (2019). Post hoc power analysis: is it an informative and meaningful analysis?. *General Psychiatry*, *32*(4), 1-4. https://doi.org/10.1136%2Fgpsych-2019-100069

Zhu, Q. & Lowe, P. (2018). Split-half reliability. In B. Frey (Eds.), *The SAGE encyclopedia of educational research, measurement, and evaluation* (Vol. 1, pp. 1573-1574). Thousand Oaks,, CA: SAGE Publications, Inc. https://doi.org/10.4135/9781506326139.n653